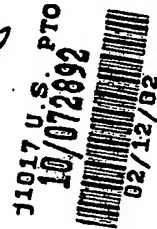


#3 4-25-02  
*Priority Paper*  
IN THE U.S. PATENT AND TRADEMARK OFFICE



Applicant(s): TASAKI, Hirohisa

Application No.:

Group:

Filed: February 12, 2002

Examiner:

For: SPEECH CODING METHOD AND SPEECH CODING APPARATUS

LETTER

Assistant Commissioner for Patents  
Box Patent Application  
Washington, D.C. 20231

February 12, 2002  
1163-0390P

Sir:

Under the provisions of 35 USC 119 and 37 CFR 1.55(a), the applicant hereby claims the right of priority based on the following application(s):

<u>Country</u>	<u>Application No.</u>	<u>Filed</u>
JAPAN	2001-52944	02/27/01

A certified copy of the above-noted application(s) is(are) attached hereto.

If necessary, the Commissioner is hereby authorized in this, concurrent, and future replies, to charge payment or credit any overpayment to deposit Account No. 02-2448 for any additional fees required under 37 C.F.R. 1.16 or under 37 C.F.R. 1.17; particularly, extension of time fees.

Respectfully submitted

BIRCH, STEWART, KOLASCH & BIRCH, LLP

By:

MICHAEL K. MUTTER

Reg. No. 29,680

P. O. Box 747

Falls Church, Virginia 22040-0747

Attachment  
(703) 205-8000  
/sll

TASAKI, Hirohisa  
February 12, 2002  
10/072,892 BSKB, LLP  
(703)205-8000  
1163-0390P  
1 of 1

日本国特許庁  
JAPAN PATENT OFFICE

紙添付の書類に記載されている事項は下記の出願書類に記載されて  
事項と同一であることを証明する。

is to certify that the annexed is a true copy of the following application as filed  
his Office.

J10170.S. PTO  
10/072892  
02/12/02

願年月日  
of Application: 2001年 2月27日

願番号  
cation Number: 特願2001-052944

願人  
ant(s): 三菱電機株式会社

2001年11月 2日

特許庁長官  
Commissioner,  
Japan Patent Office

及川耕造

出証番号 出証特 2001-3095957

【書類名】 特許願

【整理番号】 529426JP01

【提出日】 平成13年 2月27日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 9/14

【発明者】

【住所又は居所】 東京都千代田区丸の内二丁目2番3号 三菱電機株式会社  
社内

【氏名】 田崎 裕久

【特許出願人】

【識別番号】 000006013

【氏名又は名称】 三菱電機株式会社

【代理人】

【識別番号】 100066474

【弁理士】

【氏名又は名称】 田澤 博昭

【選任した代理人】

【識別番号】 100088605

【弁理士】

【氏名又は名称】 加藤 公延

【手数料の表示】

【予納台帳番号】 020640

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声符号化方法および音声符号化装置

【特許請求の範囲】

【請求項 1】 複数の音源モードの中から 1 つの音源モードを選択し、その音源モードを使用して、入力音声を所定長区間よりなるフレーム毎に符号化する音声符号化方法において、

符号化工程により、前記各音源モード毎に前記入力音声から求まる符号化対象信号の符号化を行って、その時の符号化歪を出力し、

比較工程により、前記符号化工程にて符号化された符号化歪と、固定の閾値または前記符号化対象信号の信号パワーに基づいて決定した閾値との比較を行い、

選択工程により、前記符号化工程にて符号化された符号化歪と、前記比較工程による比較結果に基づいて、前記音源モードの選択を行うことを特徴とする音声符号化方法。

【請求項 2】 複数の音源モードの中から 1 つの音源モードを選択し、その音源モードを使用して、入力音声を所定長区間よりなるフレーム毎に符号化する音声符号化方法において、

符号化工程により、前記各音源モード毎に前記入力音声から求まる符号化対象信号の符号化を行って、その時の符号化歪を出力し、

選択工程により、前記符号化工程にて符号化された符号化歪の相互比較を行って、その比較結果に基づいて音源モードの 1 つを選択し、

比較工程により、前記選択工程にて選択された音源モードに対応する符号化歪と、固定の閾値または前記符号化対象信号の信号パワーに基づいて決定した閾値との比較を行い、

置換工程により、前記比較工程による比較結果に基づいて、前記選択工程にて選択された音源モードを置換することを特徴とする音声符号化方法。

【請求項 3】 選択工程では、符号化歪が閾値を上回る比較結果が得られている音源モードの選択を抑制するようにしたことを特徴とする請求項 1 記載の音声符号化方法。

【請求項 4】 閾値を、音源モード毎に用意したことを特徴とする請求項 1



記載の音声符号化方法。

【請求項 5】 符号化歪の出力変換を行う変換工程により、比較工程によるあらかじめ定めた音源モードの符号化歪と閾値の比較の結果、前記符号化歪が前記閾値を上回る場合に、その符号化歪を前記閾値の値で置換し、

選択工程により、前記変換工程より出力された符号化歪を含む全ての音源モードの符号化歪の中から、最小の符号化歪に対応した音源モードを選択することを特徴とする請求項 1 記載の音声符号化方法。

【請求項 6】 置換工程により、選択工程が選択した音源モードに対応する符号化歪が閾値を上回る場合に、あらかじめ定めた音源モードを選択することを特徴とする請求項 2 記載の音声符号化方法。

【請求項 7】 閾値を、入力音声または符号化対象信号に対して所定の歪率となるように設定したことを特徴とする請求項 1 または請求項 2 記載の音声符号化方法。

【請求項 8】 判定工程を設けて、入力音声または符号化対象信号の分析を行って音声様態を判定し、

前記判定工程が所定の判定結果を出力したときに限って、選択工程が比較工程による比較の結果を使用せずに音源モードの選択を行うことを特徴とする請求項 1 記載の音声符号化方法。

【請求項 9】 判定工程を設けて、入力音声または符号化対象信号の分析を行って音声様態を判定し、

閾値算出工程を設けて、前記判定工程による判定結果に基づいて閾値の決定を行い、

比較工程では、前記閾値算出工程にて決定された閾値を用いて比較を行うことを特徴とする請求項 1 または請求項 2 記載の音声符号化方法。

【請求項 10】 判定工程が、少なくとも音声の立ち上がりであるか否かの判定を行うことを特徴とする請求項 8 または請求項 9 記載の音声符号化方法。

【請求項 11】 複数の音源モードを、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードで形成したことを特徴とする請求項 1 から請求項 10 のうちのいずれか 1 項記載の音声符号化方法。

【請求項 1 2】 複数の音源モードを、非雑音的な音源符号語を使用する音源モードと、雑音的な音源符号語を使用する音源モードで形成したことを特徴とする請求項 1 から請求項 1 0 のうちのいずれか 1 項記載の音声符号化方法。

【請求項 1 3】 複数の音源モードの中から 1 つの音源モードを選択し、その音源モードを使用して、入力音声をも所定長区間よりなるフレーム毎に符号化する音声符号化装置において、

前記各音源モード毎に、前記入力音声から求まる符号化対象信号の符号化を行い、その時の符号化歪を出力する符号化手段と、

前記符号化手段にて符号化された符号化歪と、固定の閾値または前記符号化対象信号の信号パワーに基づいて決定した閾値との比較を行う比較手段と、

前記符号化手段にて符号化された符号化歪と、前記比較手段による比較の結果に基づいて、前記音源モードの選択を行う選択手段とを備えたことを特徴とする音声符号化装置。

【請求項 1 4】 複数の音源モードの中から 1 つの音源モードを選択し、その音源モードを使用して、入力音声をも所定長区間よりなるフレーム毎に符号化する音声符号化装置において、

前記各音源モード毎に、前記入力音声から求まる符号化対象信号の符号化を行い、その時の符号化歪を出力する符号化手段と、

前記符号化手段にて符号化された符号化歪を相互に比較し、その比較結果に基づいて音声モードの 1 つを選択する選択手段と、

前記選択手段にて選択された音声モードに対応する符号化歪と、固定の閾値または前記符号化対象信号の信号パワーに基づいて決定した閾値との比較を行う比較手段と、

前記選択手段にて選択された音声モードの置換を、前記比較手段による比較結果に基づいて行う置換手段とを備えたことを特徴とする音声符号化装置。

【請求項 1 5】 比較手段が、符号化手段より出力された符号化歪と比較するための閾値を、符号化対象信号に対して所定の歪率となるように設定することを特徴とする請求項 1 3 または請求項 1 4 記載の音声符号化装置。

【請求項 1 6】 入力音声または符号化対象信号を分析して音声様態を判定

する判定手段を備え、

選択手段は、前記判定手段が所定の判定結果を出したときに限って、比較手段での比較結果を使用せずに音源モードの選択を行うものであることを特徴とする請求項 1 3 記載の音声符号化装置。

【請求項 1 7】 複数の音源モードを、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードで構成したことを特徴とする請求項 1 3 から請求項 1 6 のうちのいずれか 1 項記載の音声符号化装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】

この発明は、ディジタル音声信号を少ない情報量に圧縮する音声符号化方法および音声符号化装置に関するものであり、特に、音声符号化方法および音声符号化装置における音源の符号化に関するものである。

【0 0 0 2】

【従来の技術】

従来の音声符号化方法および音声符号化装置の多くは、入力音声スペクトル包絡情報と音源に分けて、フレーム単位で各々を符号化して音声符号を生成している。音源の符号化に関しては、背景雑音区間を含む様々な様態を持つ入力音声に対する符号化品質を確保するために、表現できる音源に違いを与えた複数の音源モードを用意し、その内の 1 つをフレーム毎に選択して使用する、いわゆるマルチモード符号化が検討されている。このような従来のマルチモード符号化を行う音声符号化方法および音声符号化装置としては、例えば、特開平 3 - 1 5 6 4 9 8 号公報、あるいは国際公開 W O 9 8 / 4 0 8 7 7 号公報などに開示されているものがある。

【0 0 0 3】

図 8 は特開平 3 - 1 5 6 4 9 8 号公報に開示されている従来の音声符号化装置の構成を示すブロック図である。図において、1 は入力音声、2 は線形予測分析手段、3 は線形予測係数符号化手段、7 は多重化手段、8 は音声符号、4 7 は音源符号化部である。また、音源符号化部 4 7 内において、4 8 は分類手段、4 9

、50は切換手段、51はマルチパルス音源符号化手段、52は母音部音源符号化手段である。

【0004】

次に、この特開平3-156498号公報に開示された従来の音声符号化装置の動作について説明する。

ここで、図示のように構成された従来の音声符号化装置では、あらかじめ定められた区間長、例えば10msを1フレームとしてフレーム単位で処理を行う。

【0005】

まず、入力音声1が線形予測分析手段2と分類手段48と切換手段49に入力される。線形予測分析手段2はその入力音声1を分析し、音声のスペクトル包絡情報である線形予測係数を抽出する。線形予測係数符号化手段3はこの抽出された線形予測係数を符号化し、その符号を多重化手段7に出力するとともに、音源の符号化のために量子化された線形予測係数を出力する。

【0006】

分類手段48は入力音声1の音響的特徴を分析して、母音性信号とそれ以外とに分類し、分類結果を切換手段49と切換手段50に出力する。切換手段49は分類手段48による分類結果が母音性信号である場合に、入力音声1を母音部音源符号化手段52に接続し、分類手段48による分類結果が母音性信号でない場合に、入力音声1をマルチパルス音源符号化手段51に接続する。

【0007】

マルチパルス音源符号化手段51は複数のパルス列の組合せによって音源を符号化し、符号化結果を切換手段50に出力する。母音部音源符号化手段52は可変時間長のセグメント長を算出するとともに、このセグメントの音源信号を、例えば改良ピッチ補間マルチパルス音源モデルを用いて符号化し、その符号化結果を切換手段50に出力する。

【0008】

切換手段50は、分類手段48による分類結果が母音性信号である場合に、母音部音源符号化手段52が出力した符号化結果を多重化手段7に接続し、分類手段48による分類結果が母音性信号でない場合に、マルチパルス音源符号化手段

5 1 が出力した符号化結果を多重化手段 7 に接続する。多重化手段 7 は、線形予測係数符号化手段 3 から入力された符号と、切換手段 5 0 から入力された符号化結果を多重化して、得られた音声符号 8 を出力する。

## 【 0 0 0 9 】

このように、特開平 3 - 1 5 6 4 9 8 号公報に開示された従来の音声符号化装置では、入力音声 1 の音響的特徴に基づいてあらかじめ複数種類用意した音源モデルの中から 1 つを選択し、選択した音源モデルを用いた符号化を行うことで、音声信号を少ない情報量で良好に表すことができる、と報告されている。

## 【 0 0 1 0 】

また、図 9 は、国際公開 W O 9 8 / 4 0 8 7 7 号公報に開示されている従来の音声符号化装置の構成を示すブロック図である。図において、1 は入力音声、2 は線形予測分析手段、3 は線形予測係数符号化手段、4 は適応音源符号化手段、7 は多重化手段、8 は音声符号、5 3, 5 4 は駆動音源符号化手段、5 5, 5 6 はゲイン符号化手段、5 7 は最小歪選択手段である。

## 【 0 0 1 1 】

次に、この国際公開 W O 9 8 / 4 0 8 7 7 号公報に開示された従来の音声符号化装置の動作について説明する。

ここで、図示のように構成された従来の音声符号化装置では、5 ~ 5 0 m s 程度の長さの音声を 1 フレームとして、フレーム単位で処理を行う。音源の符号化については、1 フレームを 2 分割したサブフレーム毎に処理を行う。なお、説明を分かりやすくするために以降の説明では、フレームとサブフレームを特に区別せず、単にフレームと記す。

## 【 0 0 1 2 】

まず、入力音声 1 が線形予測分析手段 2、適応音源符号化手段 4、および駆動音源符号化手段 5 3 に入力される。線形予測分析手段 2 は入力音声 1 を分析し、音声のスペクトル包絡情報である線形予測係数を抽出する。線形予測係数符号化手段 3 は、この線形予測係数を符号化し、その符号を多重化手段 7 に出力するとともに、音源の符号化のために量子化された線形予測係数を出力する。

## 【 0 0 1 3 】

適応音源符号化手段 4 には、過去の所定長の音源（信号）が適応音源符号帳として記憶されている。この適応音源符号帳に、数ビットの 2 進数値で示した適応音源符号を入力すると、その適応音源符号から繰返し周期を算出し、この繰返し周期を用いて過去の音源を周期的に繰返した時系列ベクトルを生成して出力する。適応音源符号化手段 4 はこの適応音源符号帳に各適応音源符号を入力することによって得られる各時系列ベクトルを、線形予測係数符号化手段 3 からの量子化された線形予測係数を用いた合成フィルタに通すことによって、仮の合成音を求める。そして、この仮の合成音に適切なゲインを乗算した信号と、入力音声 1 との間の歪を調べる。この処理を全ての適応音源符号に対して行い、最小の歪を与えた適応音源符号を選択するとともに、選択した適応音源符号に対応する時系列ベクトルを適応音源として出力する。また、入力音声 1 から適応音源による合成音に適切なゲインを乗算した信号を差し引いた信号を、符号化対象信号として出力する。

## 【 0 0 1 4 】

駆動音源符号化手段 5 4 には、複数の時系列ベクトルが駆動音源符号帳として記憶されている。この駆動音源符号帳は数ビットの 2 進数値で示した駆動音源符号が入力されると、その駆動音源符号に対応する位置に格納されている時系列ベクトルを読み出して出力する。駆動音源符号化手段 5 4 はこの駆動音源符号帳に各適応音源符号を入力することによって各時系列ベクトルを求め、それを線形予測係数符号化手段 3 からの量子化された線形予測係数を用いた合成フィルタに通すことによって、仮の合成音を求める。そして、この仮の合成音に適切なゲインを乗算した信号と、適応音源符号化手段 4 から入力された符号化対象信号との歪を調べる。この処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択するとともに、選択された駆動音源符号に対応する時系列ベクトルを駆動音源として出力する。

## 【 0 0 1 5 】

ゲイン符号化手段 5 6 には、適応音源と駆動音源に対する 2 つのゲイン値を表す複数のゲインベクトルがゲイン符号帳として記憶されている。このゲイン符号帳は数ビットの 2 進数値で示したゲイン符号が入力されると、そのゲイン符号に

対応する位置に格納されているゲインベクトルを読み出して出力する。ゲイン符号化手段56はこのゲイン符号帳に各ゲイン符号を入力することによってゲインベクトルを求め、その第一要素を適応音源符号化手段4から出力された適応音源に乗算し、またゲインベクトルの第二要素を駆動音源符号化手段54から出力された駆動音源に乗算し、得られた2つの信号を加算して仮の音源を生成する。そして、この仮の音源を線形予測係数符号化手段3からの量子化された線形予測係数を用いた合成フィルタに通すことによって、仮の合成音を求め、その仮の合成音と駆動音源符号化手段54を介して入力された入力音声1との歪を調べる。この処理を全てのゲイン符号に対して行い、最小の歪を与えたゲイン符号を選択する。そして、選択されたゲイン符号、適応音源符号化手段4から駆動音源符号化手段54を介して入力された適応音源符号、および駆動音源符号化手段54から入力された駆動音源符号からなる音源符号と、最小の歪と、選択されたゲイン符号に対応する仮の音源とを最小歪選択手段57に出力する。

## 【0016】

一方、駆動音源符号化手段53には、複数の時系列ベクトルが駆動音源符号帳として記憶されている。この駆動音源符号帳は数ビットの2進数値で示した駆動音源符号が入力されると、その駆動音源符号に対応する位置に格納されている時系列ベクトルを読み出して出力する。駆動音源符号化手段53はこの駆動音源符号帳に各適応音源符号を入力することによって各時系列ベクトルを求め、それを線形予測係数符号化手段3からの量子化された線形予測係数を用いた合成フィルタに通すことによって、仮の合成音を求める。そして、この仮の合成音に適切なゲインを乗算した信号と、入力音声1との歪を調べる。この処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択するとともに、選択された駆動音源符号に対応する時系列ベクトルを駆動音源として出力する。

## 【0017】

ゲイン符号化手段55には、駆動音源に対する複数個のゲイン値が第一のゲイン符号帳として記憶されている。このゲイン符号帳は数ビットの2進数値で示したゲイン符号が入力されると、そのゲイン符号に対応する位置に格納されているゲイン値を読み出して出力する。ゲイン符号化手段55はこのゲイン符号帳に各

ゲイン符号を入力することによってゲイン値を得、そのゲイン値を駆動音源符号化手段 5 3 から出力された駆動音源に乗算し、得られた信号を仮の音源とする。そして、この仮の音源を線形予測係数符号化手段 3 からの量子化された線形予測係数を用いた合成フィルタに通すことによって、仮の合成音を求め、この仮の合成音と駆動音源符号化手段 5 3 を介して入力された入力音声 1 との歪を調べる。この処理を全てのゲイン符号に対して行い、最小の歪を与えたゲイン符号を選択する。そして選択したゲイン符号と、駆動音源符号化手段 5 3 から入力された駆動音源符号とからなる音源符号と、最小の歪と、選択したゲイン符号に対応する仮の音源とを最小歪選択手段 5 7 に出力する。

## 【 0 0 1 8 】

最小歪選択手段 5 7 は、ゲイン符号化手段 5 5 から入力された最小の歪と、ゲイン符号化手段 5 6 から入力された最小の歪を比較し、より小さい歪を出力したゲイン符号化手段 5 5 または 5 6 を選択し、選択したゲイン符号化手段 5 5, 5 6 が出力した音源符号を多重化手段 7 に出力する。また、選択したゲイン符号化手段 5 5, 5 6 が出力した仮の音源を最終的な音源として適応音源符号化手段 4 に対して出力する。適応音源符号化手段 4 は、最小歪選択手段 5 7 から入力した音源を用いて、内部の適応音源符号帳の更新を行う。

## 【 0 0 1 9 】

その後、多重化手段 7 は線形予測係数符号化手段 3 から出力された線形予測係数の符号と、最小歪選択手段 5 7 から出力された音源符号とを多重化し、得られた音声符号 8 を出力する。

## 【 0 0 2 0 】

このように、国際公開 W O 9 8 / 4 0 8 7 7 号公報に開示された従来の音声符号化装置では、2つの音源モードでの符号化をその双方にて行い、小さい歪を与えた音源モードを選択することで、最もよい符号化特性を与えるモード選択ができ、符号化品質が改善する、と報告されている。

## 【 0 0 2 1 】

なお、このような音声符号化装置に関連する記載がなされた文献としては、例えば、入力音声から遅延パラメータに対応した長さの目標音声ベクトルを生成し



、適応音源探索、および駆動音源探索を行う特開平 9 - 3 1 9 3 9 6 号公報、適応音源信号のパワー情報により、駆動音源に対するゲイン量子化テーブルを、複数のゲイン量子化テーブル中より選択する特開 2 0 0 0 - 1 7 5 5 9 8 号公報などもある。

## 【 0 0 2 2 】

## 【発明が解決しようとする課題】

従来の音声符号化装置は以上のように構成されているので、それぞれ以下に述べるような課題があった。

## 【 0 0 2 3 】

特開平 3 - 1 5 6 4 9 8 号公報に開示された従来の音声符号化装置では、入力音声 1 の音響的特徴だけに基づいてあらかじめ複数種類用意した音源モデルの中から 1 つを選択してしまうため、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質、つまり音質が必ずしも最良にならないという課題がある。すなわち、入力音声 1 の音響的特徴に基づく分類では、必ず分類誤りがあるため、入力音声に適切でない音源モデルを選択してしまうことがある。また、入力音声 1 の分類が正しかったとしても、選択した音源モデルでは良好な符号化ができず、むしろ選択されなかった音源モデルで符号化した方が、音声復号化装置で復号化して得られる復号音の音質がよくなる場合もある。例えば、母音区間であっても、過渡部などの波形的な乱れが多い場合には、母音部音源符号化手段 5 2 での符号化結果が悪く、むしろマルチパルスを用いた方が変化に良好に対応できる場合がある。

## 【 0 0 2 4 】

また、国際公開 W O 9 8 / 4 0 8 7 7 号公報に開示された従来の音声符号化装置では、2 つの音源モードでの符号化をその両方でそれぞれ行い、小さい歪を与えた音源モードを選択しているため、符号化歪は最小となるものの、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質（音質）が必ずしも最良にならないという課題がある。以下、図 7 を参照しながらそれについて詳細に説明する。

## 【 0 0 2 5 】

図7 (a) は入力音声を示し、同図 (b) は雑音的な音声表現するために用意した音源モードを選択したときの復号音（音声符号を音声復号化装置で復号した結果）、同図 (c) は母音的な音声表現するために用意した音源モードを選択したときの復号音を示している。なお、図7 (a) に示した入力音声は雑音的な特徴を有する区間のものであり、図示のように、雑音的な入力音声は振幅の大きい部分と小さい部分がフレーム中に混在していることが多い。

## 【0026】

図7の場合、同図 (a) と (b) の信号の差信号のパワーとして求められる歪の値は、同図 (a) と (c) の歪よりも大きくなっている。これは図7 (a) に示す入力音声の振幅が大きい部分において、同図 (c) との差が小さいためである。しかしながら、図7 (b) と (c) を人間が聴取したときには、同図 (c) の場合にはパルス的な劣化音が聞こえてしまい、図7 (b) の方が好ましい、とされてしまう。このように歪が最小の音源モードを選択する従来の音声符号化装置では、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質（音質）では必ずしも最適の選択となっていない場合がある。

## 【0027】

この発明は、上記のような課題を解決するためになされたもので、よりよい音質を与える音源を適切に選択でき、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質、つまり音質を改善することのできる音声符号化方法および音声符号化装置を得ることを目的とする。

## 【0028】

## 【課題を解決するための手段】

この発明に係る音声符号化方法は、各音源モード毎に入力音声から求められる符号化対象信号の符号化を行って、その時に符号化された符号化歪と、固定の閾値または符号化対象信号の信号パワーに基づいて決定した閾値とを比較して、その比較結果に基づいて符号化された符号化歪を選択することにより、複数の音源モードの中から1つの音源モードを選択し、この選択された音源モードを使用して、所定長区間よりなるフレーム毎に入力音声の符号化を行うようにしたものである。

## 【 0 0 2 9 】

この発明に係る音声符号化方法は、各音源モード毎に入力音声から求められる符号化対象信号の符号化を行って音源モードの1つを選択し、選択された音源モードに対応する符号化歪と、固定の閾値あるいは符号化対象信号の信号パワーに基づいて決定した閾値との比較を行って、その比較結果に基づいて選択された音源モードの置換を行い、置換された音源モードを使用して、所定長区間よりなるフレーム毎に入力音声の符号化を行うようにしたものである。

## 【 0 0 3 0 】

この発明に係る音声符号化方法は、符号化歪の選択に際して、符号化歪が閾値を上回る比較結果が得られている音源モードについては、その選択を抑制するようにしたものである。

## 【 0 0 3 1 】

この発明に係る音声符号化方法は、音源モード毎に閾値を用意するようにしたものである。

## 【 0 0 3 2 】

この発明に係る音声符号化方法は、符号化歪の出力変換を行う際に、あらかじめ定めた音源モードの符号化歪と閾値との比較の結果、前者が後者を上回る場合に、その符号化歪を閾値の値で置換し、この出力変換された符号化歪を含む全ての音源モードの符号化歪の中から、最小の符号化歪に対応した音源モードを選択するようにしたものである。

## 【 0 0 3 3 】

この発明に係る音声符号化方法は、選択された音源モードに対応する符号化歪が閾値を上回る場合に、あらかじめ定めた音源モードを選択するようにしたものである。

## 【 0 0 3 4 】

この発明に係る音声符号化方法は、入力音声または符号化対象信号に対して所定の歪率となるように、閾値の設定を行うようにしたものである。

## 【 0 0 3 5 】

この発明に係る音声符号化方法は、入力音声または符号化対象信号の分析によ

る音声様態の判定結果が、所定のものであった場合に限って、音源モードの選択を、符号化歪と閾値との比較結果を使用せずに行うようにしたものである。

## 【 0 0 3 6 】

この発明に係る音声符号化方法は、入力音声または符号化対象信号の分析を行って音声様態を判定し、その判定結果に基づいて決定された閾値を用いて、符号化歪と閾値との比較を行うようにしたものである。

## 【 0 0 3 7 】

この発明に係る音声符号化方法は、入力音声または符号化対象信号の分析を行って音声様態を判定する際、少なくとも音声の立ち上がりであるか否かの判定を行うようにしたものである。

## 【 0 0 3 8 】

この発明に係る音声符号化方法は、複数の音源モードを、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードとで構成するようにしたものである。

## 【 0 0 3 9 】

この発明に係る音声符号化方法は、複数の音源モードを、非雑音的な音源符号語を使用する音源モードと、雑音的な音源符号語を使用する音源モードとで構成するようにしたものである。

## 【 0 0 4 0 】

この発明に係る音声符号化装置は、各音源モード毎に、入力音声から求まる符号化対象信号の符号化を行う符号化手段、符号化された符号化歪と、固定の閾値または符号化対象信号の信号パワーに基づいて決定した閾値との比較を行う比較手段、および符号化された符号化歪と、比較手段の比較結果に基づいて、音源モードの選択を行う選択手段を備え、複数の音源モードの中から選択した1つの音源モードを使用して、入力音声を所定長区間よりなるフレーム毎に符号化するようにしたものである。

## 【 0 0 4 1 】

この発明に係る音声符号化装置は、各音源モード毎に、入力音声から求まる符号化対象信号の符号化を行う符号化手段、符号化された符号化歪の相互比較結果

に基づいて音声モードの1つを選択する選択手段、選択された音声モードに対応する符号化歪と、固定の閾値または符号化対象信号の信号パワーに基づいて決定した閾値との比較を行う比較手段、および比較手段の比較結果に基づいて、選択された音声モードの置換を行う置換手段を備え、複数の音源モードの中から選択した1つの音源モードを使用して、入力音声をもとに所定長区間よりなるフレーム毎に符号化するようにしたものである。

#### 【0042】

この発明に係る音声符号化装置は、比較手段にて、符号化手段より出力された符号化歪と比較するための閾値を、符号化対象信号に対して所定の歪率となるように設定するようにしたものである。

#### 【0043】

この発明に係る音声符号化装置は、判定手段を備えて、入力音声または符号化対象信号の分析により音声様態の判定を行い、この判定手段が所定の判定結果を出力したときに限って、選択手段が、比較手段での比較結果を使用せずに音源モードの選択を行うようにしたものである。

#### 【0044】

この発明に係る音声符号化装置は、複数の音源モードを、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードとで構成するようにしたものである。

#### 【0045】

#### 【発明の実施の形態】

以下、この発明の実施の一形態を説明する。

#### 実施の形態1.

図1はこの発明の実施の形態1による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。図において、1は当該音声符号化装置に入力される入力音声であり、2はその入力音声1より線形予測係数を抽出する線形予測分析手段、3は抽出された線形予測係数を符号化するために量子化する線形予測係数符号化手段である。4は入力音声1と線形予測係数符号化手段3からの信号をもとに、適応音源と符号化対象信号を出力する適応音源符号化手段である。

5は入力音声1と、線形予測係数符号化手段3からの信号および適応音源符号化手段4からの信号をもとに、駆動音源と駆動音源符号、およびモード選択情報を出力する駆動音源符号化部である。6は入力音声1と、線形予測係数符号化手段3からの信号、および駆動音源符号化部5からの信号を入力としてゲイン符号を選択するとともに、当該ゲイン符号に対応する音源を適応音源符号化手段4に入力するゲイン符号化手段である。7は線形予測係数符号化手段3、適応音源符号化手段4、駆動音源符号化部5、およびゲイン符号化手段6からの信号を多重化する多重化手段であり、8はこの多重化手段7より出力される、当該音声符号化装置で符号化された音声符号である。

## 【0046】

また、駆動音源符号化部5内において、9は乱数で生成した時系列ベクトルによる駆動音源符号帳を備えて、線形予測係数符号化手段3と適応音源符号化手段4からの信号をもとに、仮の合成音と符号化対象信号との歪を調べて、駆動音源符号、歪、および駆動音源を出力する、符号化手段としての駆動音源符号化手段である。10、11はそれぞれ異なるパルス位置テーブルを含む駆動音源符号帳を備えて、線形予測係数符号化手段3と適応音源符号化手段4からの信号をもとに、仮の合成音と符号化対象信号との歪を調べて、駆動音源符号、歪、および駆動音源を出力する符号化手段としての駆動音源符号化手段である。12は入力音声1の信号パワーを計算するパワー計算手段であり、13はパワー計算手段12からの信号より、歪に関する閾値を計算する閾値計算手段である。14は入力音声1を分析し、音声の立ち上がり部分であるか否かの判定を行う判定手段である。15は駆動音源符号化手段9からの信号と閾値計算手段13からの閾値を比較する比較手段であり、16は判定手段14の判定結果と比較手段15の比較結果に基づいて、駆動音源符号化手段9の出力変換を行う変換手段である。17は変換手段16からの信号と、駆動音源符号化手段10および11からの信号に基づいて、駆動音源と駆動音源符号、およびモード選択情報を多重化手段7に出力する、選択手段としての最小歪選択手段である。

## 【0047】

次に動作について説明する。

この実施の形態 1 による音声符号化装置では、例えば 20ms を 1 フレームとして、フレーム単位で処理を行う。音源の符号化処理、つまり適応音源符号化手段 4、駆動音源符号化部 5 およびゲイン符号化手段 6 の処理については、1 フレームを 2 分割したサブフレーム毎に処理を行う。なお、説明を分かりやすくするため、従来の場合と同様に、以降の説明ではフレームとサブフレームを特に区別せず単にフレームと記す。

## 【0048】

まず、入力音声 1 が、線形予測分析手段 2、適応音源符号化手段 4、駆動音源符号化部 5、およびゲイン符号化手段 6 に入力される。なお、駆動音源符号化部 5 に入力された入力音声 1 は、そのパワー計算手段 12 と判定手段 14 に送られる。線形予測分析手段 2 は入力音声 1 が入力されるとその分析を行い、音声のスペクトル包絡情報である線形予測係数を抽出して、線形予測係数符号化手段 3 に送出する。線形予測係数符号化手段 3 はこの線形予測分析手段 2 から受け取った線形予測係数を符号化して多重化手段 7 に出力するとともに、音源の符号化のために量子化された線形予測係数を、適応音源符号化手段 4、駆動音源符号化部 5、およびゲイン符号化手段 6 に出力する。駆動音源符号化部 5 では、この線形予測係数符号化手段 3 からの量子化された線形予測係数が、駆動音源符号化手段 9 ~ 11 に入力される。

## 【0049】

なお、この実施の形態 1 では、スペクトル包絡情報として線形予測係数を使用しているが、これに限定されるものではなく、LSP (Line Spectrum Pairs) など他のパラメータを使用してもかまわない。

## 【0050】

適応音源符号化手段 4 には、過去の所定長の音源 (信号) が記憶された適応音源符号帳が備えられている。この適応音源符号帳は、数ビットの 2 進数値で示した適応音源符号が入力されると、その適応音源符号に対応した過去の音源の繰返し周期を求め、その繰返し周期を用いて過去の音源を周期的に繰返した時系列ベクトルを生成して出力する。適応音源符号化手段 4 はこの適応音源符号帳に各適応音源符号を入力することによって得られる各時系列ベクトルを、線形予測係

数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタを通してフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と入力音声1との差分から、それら両者間の歪を調べる。

## 【0051】

適応音源符号化手段4はこの処理を全ての適応音源符号に対して行い、最小の歪を与えた適応音源符号を選択し、選択された適応音源符号に対応する時系列ベクトルを、適応音源として駆動音源符号化手段9、および駆動音源符号化手段10、11に出力する。また入力音声1から適応音源による合成音に適切なゲインを乗算した信号を差し引いた信号（両者間の歪）を、符号化対象信号として駆動音源符号化手段9、および駆動音源符号化手段10、11に出力する。

## 【0052】

駆動音源符号化手段9には、乱数によって生成された時系列ベクトルが、雑音的な音源符号語として複数、駆動音源符号帳に記憶されている。この駆動音源符号化手段9内の駆動音源符号帳は、数ビットの2進数値で示した駆動音源符号が入力されると、その駆動音源符号に対応する位置より、そこに格納されている時系列ベクトルを読み出して出力する。このようにした場合、出力される時系列ベクトルは雑音的な音源を生成している。駆動音源符号化手段9はこの駆動音源符号帳に各駆動音源符号を入力することによって得られる時系列ベクトルを、線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタを通してフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と、適応音源符号化手段4から入力された符号化対象信号との間の差分に基づく歪を調べる。ここで、符号化対象信号を $x$ 、仮の合成音を $y$ とすると、両者間の歪 $D$ は次に示す(1)式で計算される。

## 【0053】



【数 1】

$$D = \sum_i x_i^2 - \frac{(\sum_i x_i y_i)^2}{\sum_i y_i^2} \dots\dots (1)$$

【 0 0 5 4 】

駆動音源符号化手段 9 はこの処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択するとともに、選択された駆動音源符号に対応する時系列ベクトルを駆動音源として、比較手段 1 5 および変換手段 1 6 に出力する。その時、この駆動音源とともに、上記最小の歪および駆動音源符号も比較手段 1 5 と変換手段 1 6 に出力される。

【 0 0 5 5 】

駆動音源符号化手段 1 0 には、パルス位置テーブルを含んだ駆動音源符号帳が記憶されている。この駆動音源符号化手段 1 0 内の駆動音源符号帳は、数ビットの 2 進数値で示した駆動音源符号が入力されると、その駆動音源符号を複数のパルス位置符号と極性に分離して、パルス位置テーブル中の各パルス位置符号に対応する位置に格納されているパルス位置を読み出し、そのパルス位置と極性に従って、複数本のパルスを持った時系列ベクトルを生成して出力する。つまり、出力される時系列ベクトルは、複数本のパルスから成る非雑音的な音源を生成している。なお、この駆動音源符号化手段 1 0 の駆動音源符号帳には、上記パルス位置テーブル等によって非雑音的な音源符号語が格納されていると見なすことができる。

【 0 0 5 6 】

駆動音源符号化手段 1 0 はこの駆動音源符号帳に各適応音源符号を入力することによって得られる時系列ベクトルを、適応音源符号化手段 4 で選択された適応音源符号に対応する繰返し周期を用いてピッチ周期化し、さらに線形予測係数符号化手段 3 の出力する、量子化された線形予測係数を用いた合成フィルタを通してフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と適応音源符号化手段 4 から入力された符号化対象信号との差分に基づく両者間の歪を調べる。

## 【 0 0 5 7 】

駆動音源符号化手段 1 0 はこの処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択するとともに、選択された駆動音源符号に対応する時系列ベクトルを駆動音源とする。そしてこの駆動音源を、最小の歪および駆動音源符号とともに最小歪選択手段 1 7 に出力する。

## 【 0 0 5 8 】

駆動音源符号化手段 1 1 には、駆動音源符号化手段 1 0 のそれとは異なったパルス位置テーブルを含む駆動音源符号帳が記憶されている。この駆動音源符号化手段 1 1 内の駆動音源符号帳は、数ビットの 2 進数値で示した駆動音源符号が入力されると、その駆動音源符号を複数のパルス位置符号と極性に分離して、パルス位置テーブル中の各パルス位置符号に対応する位置に格納されているパルス位置を読み出し、そのパルス位置と極性に従って、複数本のパルスを持った時系列ベクトルを生成して出力する。この場合も駆動音源符号化手段 1 0 と同様に、時系列ベクトルが複数本のパルスから成る非雑音的な音源を生成し、その駆動音源符号帳には、パルス位置テーブル等によって非雑音的な音源符号語が格納されていると見なせる。

## 【 0 0 5 9 】

駆動音源符号化手段 1 1 はこの駆動音源符号帳に各適応音源符号を入力することによって得られる時系列ベクトルを、適応音源符号化手段 4 で選択された適応音源符号に対応する繰返し周期を用いてピッチ周期化し、さらに線形予測係数符号化手段 3 の出力する、量子化された線形予測係数を用いた合成フィルタを通してフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と適応音源符号化手段 4 から入力された符号化対象信号との差分をとり、それに基づく両者間の歪を調べる。

## 【 0 0 6 0 】

駆動音源符号化手段 1 1 はこの処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択するとともに、選択された駆動音源符号に対応する時系列ベクトルを駆動音源とする。そしてこの駆動音源を、最小の歪および駆動音源符号とともに最小歪選択手段 1 7 に出力する。

## 【 0 0 6 1 】

パワー計算手段 1 2 は受け取った入力音声 1 のフレーム内の信号パワーを計算し、得られた信号パワーを閾値計算手段 1 3 に出力する。閾値計算手段 1 3 はこのパワー計算手段 1 2 から入力された信号パワーに、あらかじめ用意されている歪率に関する定数を乗算し、その計算結果を歪に関する閾値として比較手段 1 5 と変換手段 1 6 に出力する。

## 【 0 0 6 2 】

ここで、あらかじめ用意した定数を  $R$ 、信号パワーを  $P$  とすると、歪に関する閾値  $D_{th}$  は、次の (2) 式で求めることができる。

$$D_{th} = R \cdot P \quad \dots (2)$$

## 【 0 0 6 3 】

なお、定数  $R$  はパワー領域での歪率に関する値であり、この実施の形態 1 では 0.7 とする。また、入力音声 1 の信号パワー  $P$  にこの歪率に関する定数  $R$  を乗算して得られる、歪に関する閾値  $D_{th}$  は、(1) 式に示した歪の領域で定義される値となる。

## 【 0 0 6 4 】

一方、判定手段 1 4 は受け取った入力音声 1 の分析を行って音声態様を判定する。その結果、音声の立ち上がり部分である場合には“0”を、それ以外の場合には“1”を、それぞれ判定結果として出力する。音声の立ち上がりであるか否かは、入力音声 1 の信号パワーを前フレームの信号パワーで除した結果が、所定の閾値を超えているか否かで大まかに判定することができる。

## 【 0 0 6 5 】

比較手段 1 5 は駆動音源符号化手段 9 から入力された歪  $D$  と、閾値計算手段 1 3 から入力された歪に関する閾値  $D_{th}$  との比較を行い、その比較結果として、歪  $D$  の方が大きいときには“1”を、それ以外の場合には“0”をそれぞれ出力する。変換手段 1 6 は判定手段 1 4 から出力された判定結果と比較手段 1 5 から出力された比較結果とを受け、その両方が“1”であった場合には、駆動音源符号化手段 9 から出力された歪  $D$  を、閾値計算手段 1 3 から入力された閾値  $D_{th}$  の値に置換する。なお、この変換手段 1 6 は判定手段 1 4 の判定結果、あるいは比較

手段15の比較結果のいずれか一方でも“0”であった場合には、上記置換の処理は行わない。この変換手段16による置換処理結果は最小歪選択手段17に出力される。

## 【0066】

最小歪選択手段17はこの変換手段16から入力された歪と、駆動音源符号化手段10から入力された歪と、駆動音源符号化手段11から入力された歪との比較を行って、これらの中で最小の歪を選択する。そして、この選択された歪を出力した変換手段16または駆動音源符号化手段10、11より出力された駆動音源をゲイン符号化手段6に、駆動音源符号を多重化手段7にそれぞれ出力する。さらに、これら3つの歪の内のどれを選択したかを示す情報をモード選択情報として多重化手段7に出力する。

## 【0067】

なお、(1)式の第一項は仮の合成音 $y$ に依存しないので、歪 $D$ を最小化する $y$ を探索することは、次の(3)式に示した、上記(1)式の第二項を最大化する $y$ を探索することと等価である。

## 【0068】

【数2】

$$d = \frac{(\sum_i x_i y_i)^2}{\sum_i y_i^2} \quad \dots\dots (3)$$

## 【0069】

従って、複数の仮の合成音 $y$ に対して、この(3)式で示した評価値 $d$ を計算し、これを最大化する仮の合成音 $y$ を与えた駆動音源符号を選択しても同じ結果となる。ただし、各駆動音源符号化手段が(3)式で示した評価値 $d$ を最大化する駆動音源符号を探索し、歪 $D$ の代わりに評価値 $d$ を出力する場合には、これに応じて閾値計算手段13、比較手段15、変換手段16、最小歪選択手段17での処理を以下の通り変更する必要がある。

## 【0070】

すなわち、閾値計算手段13では、符号化対象信号 $x$ の信号パワーを $P'$ とし

て、評価値  $d$  に対応する閾値  $d_{th}$  を次の (4) 式により算出する。

$$d_{th} = P' - R \cdot P \quad \dots (4)$$

【0071】

ここで、この (4) 式は、(1) 式と (3) 式とを合わせて次の (5) 式を求め、得られた (5) 式の第二項に (2) 式を代入することで導出できる。なお、

(5) 式の第一項は符号化対象信号の信号パワー  $P'$  である。この時、閾値計算手段 13 に対する入力として、適応音源符号化手段 4 から出力された符号化対象信号を追加する必要がある。

【0072】

【数 3】

$$d_{th} = \sum_i x_i^2 - D_{th} \quad \dots (5)$$

【0073】

また、比較手段 15 では、駆動音源符号化手段 9 が出力した評価値  $d$  と、閾値計算手段 13 から入力された閾値  $d_{th}$  との比較を行い、その比較結果として、評価値  $d$  の方が小さい場合には “1” を、それ以外の場合には “0” をそれぞれ出力する。変換手段 16 にはこの比較手段 15 が出力した比較結果と、判定手段 14 が出力した判定結果が入力され、その両方がともに “1” であれば、駆動音源符号化手段 9 から出力された結果中の評価値  $d$  を、閾値計算手段 13 から入力された閾値  $d_{th}$  の値に置換する。なお、それ以外の場合には評価値  $d$  の置換処理は行わない。

【0074】

最小歪選択手段 17 には変換手段 16 と、駆動音源符号化手段 10 および 11 より評価値  $d$  が入力されている。最小歪選択手段 17 はそれら 3 つの評価値  $d$  の比較を行って、その中で最大の評価値を選択する。そして選択された評価値を出力した変換手段 16 または駆動音源符号化手段 10 または駆動音源符号化手段 11 が出力した駆動音源をゲイン符号化手段 6 に、駆動音源符号を多重化手段 7 にそれぞれ出力する。最小歪選択手段 17 はさらに、上記 3 つの評価値の内のどれを選択したかを示す情報をモード選択情報として多重化手段 7 に出力する。

## 【 0 0 7 5 】

ゲイン符号化手段 6 には、適応音源と駆動音源に対する 2 つのゲイン値を表す複数のゲインベクトルがゲイン符号帳として記憶されている。このゲイン符号帳は、数ビットの 2 進数値で示したゲイン符号が入力されると、そのゲイン符号に対応する位置に格納されているゲインベクトルを読み出して出力する。ゲイン符号化手段 6 はこのゲイン符号帳に各ゲイン符号を入力することによってゲインベクトルを求め、その第一要素を適応音源符号化手段 4 の出力する適応音源に乗算するとともに、第二要素を最小歪選択手段 1 7 の出力する駆動音源に乗算して、得られた 2 つの信号を加算することにより仮の音源を生成する。そしてこの仮の音源を線形予測係数符号化手段 3 の出力する、量子化された線形予測係数を用いた合成フィルタを通してフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音と入力音声 1 との差分をとり、それに基づく両者間の歪を調べる。

## 【 0 0 7 6 】

ゲイン符号化手段 6 はこの処理を全ての駆動音源符号に対して行い、最小の歪を与えたゲイン符号を選択するとともに、その選択されたゲイン符号を多重化手段 7 に、選択されたゲイン符号に対応する仮の音源を最終的な音源として適応音源符号化手段 4 にそれぞれ出力する。

## 【 0 0 7 7 】

適応音源符号化手段 4 は、このゲイン符号化手段 6 より出力された最終的な音源を受け取ると、内部に記憶している適応音源符号帳をその最終的な音源に基づいて更新する。

## 【 0 0 7 8 】

その後、多重化手段 7 は、線形予測係数符号化手段 3 から出力された線形予測係数の符号、適応音源符号化手段 4 から出力された適応音源符号、駆動音源符号化部 5 中の最小歪選択手段 1 7 から出力された駆動音源符号とモード選択情報、およびゲイン符号化手段 6 から出力されたゲイン符号を多重化し、得られた音声符号 8 を出力する。

## 【 0 0 7 9 】

次に、この実施の形態 1 により得られた音声符号 8 を音声復号化装置で復号化して得られる復号音の主観的な品質、つまり音質の改善がはかれることを、図 7 を参照しながら説明する。図 7 は符号化歪を最小にする音源モードの選択について説明するための各波形を示すイメージ図であり、図 7 (a) は入力音声、同図 (b) は雑音的な音声を表現するために用意した音源モードを選択したときの復号音（音声符号を音声復号化装置で復号した結果）、同図 (c) は母音的な音声を表現するために用意した音源モードを選択したときの復号音をそれぞれ示している。なお、図 7 (a) に示した入力音声は雑音的な特徴を有する区間のものであり、図示のように、雑音的な入力音声は振幅の大きい部分と小さい部分がフレーム中に混在していることが多い。

#### 【0080】

入力音声 1 が図 7 (a) に示すように雑音的である場合には、一般にモデル化がうまく働かないので、同図 (b) に示した雑音的な音声を表現するために用意した音源モード（雑音的な音源符号語を使用する音源モード）の場合でも、同図 (c) に示した母音的な音声を表現するために用意した音源モード（非雑音的な音源符号語を使用する音源モード）の場合でも、符号化時の歪率は比較的大きな値となる。

#### 【0081】

ここで、駆動音源符号化手段 9 は乱数によって生成された時系列ベクトルを使用しており、図 7 (b) に示す雑音的な音声を表現するために用意した音源モードに対応する。また駆動音源符号化手段 10 および 11 はパルス音源とピッチ周期化を使用しており、図 7 (c) に示す母音的な音声を表現するために用意した音源モードに対応する。

#### 【0082】

前述のように、各駆動音源符号化手段 9 ～ 11 から出力される歪  $D$  は何れも大きい値となっているが、駆動音源符号化手段 9 から出力された歪  $D$  だけは、変換手段 16 によって、歪  $D$  よりも小さい閾値  $D_{th}$  に置換される。その結果、最小歪選択手段 17 において、駆動音源符号化手段 9 が出力した駆動音源符号が選択され、復号音は図 7 (b) に示すものとなる。このように、図 7 (b) に示す復号

音の歪の方が同図(c)に示す復号音の歪よりも大きな場合でも、雑音的な区間などの符号化時の歪率が大きくなる区間では、安定的に図7(b)に示す復号音が選択される。

## 【0083】

なお、この実施の形態1では、判定手段14が音声の立ち上がり以外であると判定したときに限って、変換手段16による置換の処理を行うようにしている。すなわち、音声の立ち上がりと判定された場合でも変換手段16による置換の処理を行うようにして、復号音が図7(b)となるようにしてしまうと、破裂音のパルス的な特徴がつぶれてしまったり、母音の立ち上がりをザラザラした音質に劣化させてしまうことになる。

## 【0084】

また、この実施の形態1では、パワー計算手段12で入力音声1の信号パワーを計算し、閾値計算手段13はその信号パワーを用いて閾値の計算を行っている。すなわち、入力音声1の信号パワーに歪率に関する定数を乗算することで、一定の歪率(SN比など)となる歪の値が閾値として算出されている。この閾値を使用することで、駆動音源符号化手段9の歪が一定の歪率(SN比など)を超えたときにその歪の値を置換して、駆動音源符号化手段9の出力する歪が選択されやすくなるようにしている。

## 【0085】

なお、閾値計算手段13については、入力音声1の信号パワーを使用せず、固定閾値Rをそのまま出力する構成に変形することも可能である。その場合、各駆動音源符号化手段9~11が出力する歪を、入力音声1の信号パワーPで除算した値、つまり歪率として出力するように変形することにより、異なる構成によっても上述した実施の形態の場合と同じ結果を与えることができる。

## 【0086】

また、この実施の形態1では、パワー計算手段12が入力音声1の信号パワーを計算しているが、適応音源符号化手段4が出力した符号化対象信号の信号パワーを計算するように変更するようによい。その場合、閾値計算手段13が出力する閾値が、上記した入力音声1に対する歪に関する閾値ではなく、符号化



対象信号に対する歪に関する閾値となる。

【0087】

なお、定常母音区間では適応音源による符号化が良好で、符号化対象信号が入力音声に比べて低振幅で雑音的になる場合がある。上記のように、パワー計算手段12が符号化対象信号の信号パワーを計算する構成とした場合には、閾値も小さくなり、変換手段16での歪の置換が起こりやすくなる。しかしながら、定常母音区間では置換を行わずに歪を最小にする駆動音源符号化手段9から11を選択する方が望ましいので、置換を停止するために判定手段14における判定処理を修正する必要がある。具体的には、判定手段14が、音声の立ち上がり、または母音区間を検知した時に判定結果として“0”を出力し、それ以外の時に判定結果として“1”を出力するようにすればよい。母音区間の検知は、入力音声1のピッチ周期性の大きさ、適応音源符号化手段4での符号化処理中の中間パラメータなどを用いて行うことができる。

【0088】

また、この実施の形態1では、パワー計算手段12が入力音声1の信号パワーを計算し、その信号パワーを用いて閾値計算手段13が閾値の計算を行っているが、信号パワーの代わりに、振幅や、対数パワーなどを用いても、閾値計算手段13の計算式を変形することで同じ結果を得ることができる。

【0089】

また、この実施の形態1では、雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段9を1つ、また非雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段10、11の2つを備える構成としたが、前者を2つ以上としてもかまわないし、後者を1つまたは3つ以上としてもかまわないことはいうまでもない。

【0090】

また、この実施の形態1では、閾値 $D_{th}$ と歪 $D$ の比較結果に基づいて歪 $D$ を閾値 $D_{th}$ に置換するようにしているが、閾値 $D_{th}$ と歪 $D$ を入力変数とする関数を用意し、その出力値を新たな歪 $D$ と置換するようにしてもかまわない。

【0091】

また、この実施の形態1では、単純に信号間の2乗距離を歪としているが、音声符号化装置で多く使用されている聴覚重み付け歪とするようにしても当然かわない。

#### 【0092】

以上のように、この実施の形態1によれば、複数の音源モードの中から1つを選択し、この音源モードを使用して入力音声1をフレームと呼ばれる所定長の区間毎に符号化する際に、各音源モード毎に入力音声から求まる符号化対象信号の符号化を行い、その時の符号化歪と固定の閾値、あるいは符号化対象信号の信号パワーに基づいて決定した閾値との比較を行って、その比較結果に基づいて音源モードの選択を行うようにしたので、符号化歪が大きい場合でも復号音の品質劣化が少ない音源モードを選択することが可能となるため、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質、つまり音質を改善することができるという効果が得られる。

#### 【0093】

さらに、この実施の形態1によれば、あらかじめ定めた音源モードについて符号化歪と閾値の比較を行い、符号化歪が閾値を上回る場合にその符号化歪を閾値の値に置換し、全ての音源モードの符号化歪の中で最小の符号化歪に対応した音源モードを選択するようにしたので、符号化歪が大きい場合に符号化歪が置換された音源モードが選択されやすくなるため、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

#### 【0094】

さらに、この実施の形態1によれば、入力音声または符号化対象信号に対して所定の歪率となるように、閾値の設定を行うようにしたので、符号化時の歪率が所定以上である場合に、復号音の品質劣化が少ない音源モードを選択することが可能となるため、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

## 【 0 0 9 5 】

さらに、この実施の形態 1 によれば、入力音声または符号化対象信号を分析して音声様態を判定し、所定の判定結果となったときに限って、符号化歪と閾値との比較結果を使用せずに音源モードの選択を行うようにしたので、符号化歪が大きくても復号音の品質劣化を起しにくい入力音声に対しては、従来の場合と同様の音源モード選択が行われ、一層丁寧な音源モード選択となり、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

## 【 0 0 9 6 】

さらに、この実施の形態 1 によれば、音声様態の判定において、少なくとも音声の立ち上がりであるか否かを判定するようにしたので、音声の立ち上がりという符号化歪が大きくなりがちな区間とそれ以外の区間とで、符号化歪に基づく音源モード選択の制御を変えることが可能となるため、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することができて、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善でき、また、上記音声の立ち上がりの区間は、破裂音などのように雑音的な音源よりパルス的な音源の方が適するケースもあるため、符号化歪が大きくても、特定の音源モードを優先的に選択するような制御は劣化を引き起こすことが出てくるが、音声の立ち上がりの判定によってそれを回避できるなどの効果が得られる。

## 【 0 0 9 7 】

さらに、この実施の形態 1 によれば、複数の音源モードを、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードで構成したので、符号化歪が大きい場合に雑音的な音源を生成する音源モードを選択しやすくすることが可能となるため、非雑音的な音源を生成する音源モードを選択することによる劣化を回避し、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

## 【 0 0 9 8 】

さらに、この実施の形態 1 によれば、複数の音源モードを、非雑音的な音源符号語を使用する音源モードと、雑音的な音源符号語を使用する音源モードで構成

したので、符号化歪が大きい場合に雑音的な音源符号語を使用する音源モードを選択しやすくすることが可能となるため、非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避し、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

#### 【0099】

実施の形態 2.

図 2 はこの発明の実施の形態 2 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。図において、1 は入力音声、2 は線形予測分析手段、3 は線形予測係数符号化手段、6 はゲイン符号化手段、7 は多重化手段、8 は音声符号であり、これらは図 1 に同一符号を付して示した実施の形態 1 の各部と同等の部分である。

#### 【0100】

また、18 は入力音声 1 と線形予測係数符号化手段 3 からの信号をもとに、適応音源、駆動音源、音源符号、およびモード選択情報を出力する音源符号化部である。

#### 【0101】

この音源符号化部 18 内において、19 は乱数で生成した時系列ベクトルによる駆動音源符号帳を備え、入力音声 1 と線形予測係数符号化手段 3 からの信号をもとに、仮の合成音と入力音声 1 との歪を調べて、音源符号、歪、および駆動音源を出力する、符号化手段としての音源符号化手段である。20 はパルス位置テーブルを含む駆動音源符号帳を備え、入力音声 1 と線形予測係数符号化手段 3 からの信号をもとに、仮の合成音と入力音声 1 との歪を調べて、音源符号、歪、および駆動音源を出力する、符号化手段としての音源符号化手段である。21 は適応音源符号帳を備えた適応音源符号化手段と、駆動音源符号帳を備えた駆動音源符号化手段とによって構成され、入力音声 1 と線形予測係数符号化手段 3 からの信号をもとに、音源符号、歪、適応音源、および駆動音源を出力する、符号化手段としての音源符号化手段である。

#### 【0102】

22は入力音声の信号パワーを計算するパワー計算手段、23はパワー計算手段22からの信号より、歪に関する閾値を計算する閾値計算手段であり、24は入力音声1を分析し、音声の立ち上がり部分であるか否かを判定する判定手段である。25は音源符号化手段19からの信号と閾値計算手段23からの閾値を比較する比較手段である。26は判定手段24の判定結果と比較手段25の比較結果に基づいて音源符号化手段19の出力変換を行う変換手段である。27は変換手段26からの信号と、音源符号化手段20および21からの信号に基づいて、適応音源および駆動音源をゲイン符号化手段6に、音源符号およびモード選択情報を多重化手段7に出力する、選択手段としての最小歪選択手段である。

## 【0103】

このように、上記実施の形態2では、複数の音源符号化手段19～21の中から1つを選択する構成としている点で、複数の駆動音源符号化手段9～11の中の1つを選択する構成の実施の形態1とは異なっている。つまり、駆動音源符号化手段の他に適応音源符号化手段も含めた上位の音源符号化手段19～21の選択にこの発明を適用したものとなっている。

## 【0104】

次に動作について説明する。なお、ここでは、上記実施の形態1とは異なる部分を中心に、図2に基づいて説明する。

まず、入力音声1が線形予測分析手段2、ゲイン符号化手段6、および音源符号化部18に入力される。線形予測分析手段2は入力音声1が入力されるとその分析を行い、音声のスペクトル包絡情報である線形予測係数を抽出して、線形予測係数符号化手段3に送出する。線形予測係数符号化手段3はこの線形予測分析手段2から受け取った線形予測係数を符号化して多重化手段7に出力するとともに、音源の符号化のために量子化された線形予測係数を、音源符号化部18およびゲイン符号化手段6に出力する。なお、音源符号化部18では、入力音声1は音源符号化手段19～21、およびパワー計算手段22、判定手段24に入力され、線形予測係数符号化手段3からの量子化された線形予測係数は、音源符号化手段19～21に入力される。

## 【0105】

音源符号化手段19には、乱数によって生成された時系列ベクトルが、雑音的な音源符号語として複数、駆動音源符号帳に記憶されている。この音源符号化手段19内の駆動音源符号帳は、数ビットの2進数値で示した音源符号が入力されると、その音源符号に対応する位置より、そこに格納されている時系列ベクトルを読み出して出力する。なお、この出力された時系列ベクトルは雑音的な音源を生成している。音源符号化手段19はこの駆動音源符号帳に各音源符号を入力することによって得られる時系列ベクトルを、線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と入力音声1との差分をとり、それに基づく両者間の歪を調べる。

## 【0106】

音源符号化手段19はこの処理を全ての音源符号に対して行い、最小の歪を与えた音源符号を選択するとともに、選択された音源符号に対応する時系列ベクトルを駆動音源とする。そしてこの駆動音源を、上記最小の歪および音源符号とともに、比較手段25と変換手段26とに出力する。

## 【0107】

音源符号化手段20には、パルス位置テーブルを含んだ駆動音源符号帳が記憶されている。この音源符号化手段20内の駆動音源符号帳は、数ビットの2進数値で示した音源符号が入力されると、その音源符号を複数のパルス位置符号と極性に分離して、パルス位置テーブル中の各パルス位置符号に対応する位置に格納されているパルス位置を読み出し、そのパルス位置と極性に従って、複数本のパルスを持つ時系列ベクトルを生成して出力する。この時系列ベクトルは複数本のパルスによる非雑音的な音源を生成し、その駆動音源符号帳はパルス位置テーブル等により非雑音的な音源符号語が格納されていると見なせる。

## 【0108】

音源符号化手段20はこの駆動音源符号帳に各音源符号を入力することによって得られる各時系列ベクトルを、線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信

号と入力音声1との差分をとり、それに基づく両者間の歪を調べる。

【0109】

音源符号化手段20はこの処理を全ての音源符号に対して行い、最小の歪を与えた音源符号を選択するとともに、選択された音源符号に対応する時系列ベクトルを駆動音源とする。そしてこの駆動音源を、上記最小の歪および音源符号とともに、最小歪選択手段27に出力する。

【0110】

音源符号化手段21は、過去の所定長の音源（信号）が適応音源符号帳として記憶されている適応音源符号化手段と、パルス位置テーブルを含んだ駆動音源符号帳が記憶されている駆動音源符号化手段とによって構成されている。この音源符号化手段21中の適応音源符号化手段が持つ適応音源符号帳は、数ビットの2進数値で示した適応音源符号が入力されると、その適応音源符号から繰返し周期を算出し、この繰返し周期を用いて過去の音源を周期的に繰り返した時系列ベクトルを生成して出力する。また、この音源符号化手段21中の駆動音源符号化手段が持つ駆動音源符号帳は、数ビットの2進数値で示した駆動音源符号が入力されると、その駆動音源符号に対応する位置に格納されている時系列ベクトルを読み出して出力する。なお、この時系列ベクトルは複数本のパルスによる非雑音的な音源を生成し、その駆動音源符号帳はパルス位置テーブル等により非雑音的な音源符号語が格納されていると見なせる。

【0111】

音源符号化手段21の適応音源符号化手段はその適応音源符号帳に適応音源符号を入力することによって得られる各時系列ベクトルを、線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と入力音声1との差分をとり、それに基づく両者間の歪を調べる。音源符号化手段21の適応音源符号化手段はこの処理を全ての音源符号に対して行い、最小の歪を与えた適応音源符号を選択するとともに、選択された適応音源符号に対応する時系列ベクトルを適応音源として出力する。また、入力音声1と適応音源による合成音に適切なゲインを乗算した信号との差分を計

算し、それを符号化対象信号として出力する。

#### 【0112】

また、音源符号化手段21の駆動音源符号化手段はその駆動音源符号帳に駆動音源符号を入力することによって得られる各時系列ベクトルを、上記音源符号化手段21の適応音源符号化手段で選択された適応音源符号に対応する繰返し周期を用いてピッチ周期化し、さらに線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と、適応音源符号化手段から入力された符号化対象信号との差分をとり、それに基づく両者間の歪を調べる。音源符号化手段21の駆動音源符号化手段はこの処理を全ての駆動音源符号に対して行い、最小の歪を与えた駆動音源符号を選択して、選択された駆動音源符号に対応する時系列ベクトルを駆動音源とし、その駆動音源を上記最小の歪および駆動音源符号とともに出力する。

#### 【0113】

音源符号化手段21は最後に、その適応音源符号と駆動音源符号との多重化を行い、得られた結果を音源符号として、上記適応音源、駆動音源とともに最小歪選択手段27に出力する。

#### 【0114】

パワー計算手段22は受け取った入力音声1のフレーム内の信号パワーを計算し、得られた信号パワーを閾値計算手段23に出力する。閾値計算手段23はこのパワー計算手段22から入力された信号パワーに、あらかじめ用意されている歪率に関する定数を乗算し、その計算結果を歪に関する閾値として比較手段25と変換手段26に出力する。判定手段24は受け取った入力音声1の分析を行って音声態様を判定する。その結果、音声の立ち上がり部分である場合には“0”を、それ以外の場合には“1”を、それぞれ判定結果として出力する。

#### 【0115】

比較手段25は音源符号化手段19から入力された歪と、閾値計算手段23から入力された歪に関する閾値との比較を行い、その比較結果として、歪の方が大きいときには“1”を、それ以外の場合には“0”をそれぞれ出力する。変換手



段 2 6 は判定手段 2 4 から出力された判定結果と比較手段 2 5 から出力された比較結果とを受け、その両方が“1”であった場合には、音源符号化手段 1 9 から出力された歪を、閾値計算手段 2 3 から入力された閾値の値に置換する。なお、この変換手段 2 6 は判定手段 2 4 の判定結果、あるいは比較手段 2 5 の比較結果のいずれか一方でも“0”であった場合には、上記置換の処理は行わない。この変換手段 2 6 による置換処理結果は最小歪選択手段 2 7 に出力される。

## 【0 1 1 6】

最小歪選択手段 2 7 はこの変換手段 2 6 から入力された歪と、音源符号化手段 2 0 から入力された歪と、音源符号化手段 2 1 から入力された歪との比較を行って、それらの中より最小の歪を選択する。ここで、変換手段 2 6 から入力された歪を選択した場合には、ゲイン符号化手段 6 に、適応音源としての全要素がゼロの値を持つ信号と、変換手段 2 6 から入力された駆動音源とを出力し、多重化手段 7 に変換手段 2 6 から入力された音源符号を出力する。また、音源符号化手段 2 0 から入力された歪を選択した場合には、ゲイン符号化手段 6 に適応音源としての全要素がゼロの値を持つ信号と、音源符号化手段 2 0 から入力された駆動音源とを出力し、多重化手段 7 に音源符号化手段 2 0 から入力された音源符号を出力する。また、音源符号化手段 2 1 から入力された歪を選択した場合には、ゲイン符号化手段 6 に音源符号化手段 2 1 から入力された適応音源と駆動音源を、多重化手段 7 に音源符号化手段 2 1 から入力された音源符号を出力する。さらに、これら 3 つの歪の内のどれを選択したかを示す情報をモード選択情報として多重化手段 7 に出力する。

## 【0 1 1 7】

ゲイン符号化手段 6 には、適応音源と駆動音源に対する 2 つのゲイン値を表す複数のゲインベクトルがゲイン符号帳として記憶されている。このゲイン符号帳は、数ビットの 2 進数値で示したゲイン符号が入力されると、そのゲイン符号に対応する位置に格納されているゲインベクトルを読み出して出力する。ゲイン符号化手段 6 はこのゲイン符号帳に各ゲイン符号を入力することによってゲインベクトルを求め、その第一要素を音源符号化部 1 8 の出力する適応音源に乗算するとともに、第二要素を音源符号化部 1 8 の出力する駆動音源に乗算して、得られ

た 2 つの信号を加算することにより仮の音源を生成する。そしてこの仮の音源を線形予測係数符号化手段 3 の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音と入力音声 1 との差分をとり、それに基づく両者間の歪を調べる。

## 【 0 1 1 8 】

ゲイン符号化手段 6 はこの処理を全てのゲイン符号に対して行い、最小の歪を与えたゲイン符号を選択する。そして、この選択されたゲイン符号を多重化手段 7 に出力し、またこの選択されたゲイン符号に対応する仮の音源を、最終的な音源として音源符号化手段 2 1 内の適応音源符号化手段に出力する。

## 【 0 1 1 9 】

音源符号化手段 2 1 内の適応音源符号化手段は、このゲイン符号化手段 6 より出力された最終的な音源を受け取ると、内部に記憶している適応音源符号帳をその最終的な音源に基づいて更新する。

## 【 0 1 2 0 】

その後、多重化手段 7 は、線形予測係数符号化手段 3 から出力された線形予測係数の符号、音源符号化部 1 8 から出力された音源符号とモード選択情報、およびゲイン符号化手段 6 から出力されたゲイン符号を多重化し、得られた音声符号 8 を出力する。

## 【 0 1 2 1 】

なお、この発明の実施の形態 2 として、適応音源符号化手段も含めた上位の音源符号化手段を複数備えて、その中の 1 つを選択する、図 2 に示したものについて説明したが、音声符号化装置が複数の駆動音源符号化手段を備えてその中の 1 つを選択するように構成した、上記実施の形態 1 の音声符号化装置と同様の様々な変形が可能である。

## 【 0 1 2 2 】

以上のように、この実施の形態 2 によれば、適応音源符号化手段を含む上位の音源符号化手段を複数備え、その中の 1 つを選択するようにしているので、その音源符号化手段の選択においても、上記実施の形態 1 の場合と同様の効果が得ら

れる。

### 【 0 1 2 3 】

実施の形態 3.

図 3 はこの発明の実施の形態 3 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図であり、図中、図 1 の各部分と同等の部分については同一番号を付してその説明を省略する。図において、28 は入力音声 1 と、線形予測係数符号化手段 3 からの信号および適応音源符号化手段 4 からの信号をもとに、駆動音源と駆動音源符号、およびモード選択情報を出力する駆動音源符号化部である。

### 【 0 1 2 4 】

また、29 はパワー計算手段 12 からの信号より、歪に関する第一の閾値、および第二の閾値を計算する閾値計算手段である。30 は駆動音源符号化手段 10 からの信号と第一の閾値とを比較する比較手段であり、31 はこの比較手段 30 と判定手段 14 の判定結果に基づいて、駆動音源符号化手段 10 の出力を補正する変換手段としての補正手段である。32 は駆動音源符号化手段 11 からの信号と第二の閾値とを比較する比較手段であり、33 はこの比較手段 32 と判定手段 14 の判定結果に基づいて、駆動音源符号化手段 11 の出力を補正する変換手段としての補正手段である。なお、上記駆動音源符号化部 28 は、閾値計算手段 29、比較手段 30、32、補正手段 31、33 と、駆動音源符号化手段 9、10、11、パワー計算手段 12、判定手段 14、および最小歪選択手段 17 とによって構成されている。

### 【 0 1 2 5 】

次に動作について説明する。なお、ここでは、上記実施の形態 1 とは異なる部分を中心に、図 3 に基づいて説明する。

この場合も、線形予測係数符号化手段 3 で量子化された線形予測係数と、適応音源符号化手段 4 からの符号化対象信号とが、駆動音源符号化部 28 内の駆動音源符号化手段 9 ～ 11 に入力されている。この駆動音源符号化手段 9 には、乱数によって生成された複数の時系列ベクトルが駆動音源符号帳として記憶されている。駆動音源符号化手段 9 は実施の形態 1 の場合と同様に、その駆動音源符号帳

を用いて適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、最小歪選択手段 1 7 に出力する。

## 【 0 1 2 6 】

また駆動音源符号化手段 1 0 には、パルス位置テーブルを含む駆動音源符号帳が記憶されている。駆動音源符号化手段 1 0 はその駆動音源符号帳を用いて、実施の形態 1 の場合と同様に、適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、比較手段 3 0 と補正手段 3 1 とに出力する。同様に、駆動音源符号化手段 1 1 には、上記駆動音源符号化手段 1 0 のものとは異なるパルス位置テーブルを含む駆動音源符号帳が記憶されている。駆動音源符号化手段 1 1 はその駆動音源符号帳を用いて、適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、比較手段 3 2 と補正手段 3 3 とに出力する。

## 【 0 1 2 7 】

なお、この場合も、駆動音源符号化手段 9 の駆動音源符号帳には、乱数によって生成された雑音的な音源符号語が、駆動音源符号化手段 1 0, 1 1 の駆動音源符号帳には、パルス位置テーブル等による非雑音的な音源符号語がそれぞれ格納されており、また、駆動音源符号化手段 9 より出力される時系列ベクトルは、雑音的な音源を生成し、駆動音源符号化手段 1 0, 1 1 より出力される時系列ベクトルは、非雑音的な音源を生成している。

## 【 0 1 2 8 】

一方、閾値計算手段 2 9 はパワー計算手段 1 2 によって算出された信号パワーに、あらかじめ用意された歪率に関する第一の定数を乗算して、歪に関する第一の閾値を求め、あらかじめ用意された歪率に関する第二の定数を乗算して、歪に関する第二の閾値を求める。得られた歪に関する第一の閾値は比較手段 3 0 と補

正手段 3 1 に、歪に関する第二の閾値は比較手段 3 2 と補正手段 3 3 にそれぞれ出力される。なお、ここであらかじめ用意される第一および第二の歪率に関する定数については、符号化歪が大きいときに、駆動音源符号化手段 1 0 と 1 1 の中の、復号音の劣化がより大きい方の定数を小さく設定しておく。この歪率に関する定数が小さい程、小さい符号化歪において先に、後述する比較手段 3 0, 3 2 での比較結果が“1”となる。

## 【0 1 2 9】

判定手段 1 4 は実施の形態 1 の場合と同様に、入力音声 1 を分析して音声態様を判定する。判定の結果、音声の立ち上がり部分である場合には“0”を、それ以外の場合に“1”を出力する。

## 【0 1 3 0】

比較手段 3 0 は、駆動音源符号化手段 1 0 から入力された歪と、閾値計算手段 2 9 から入力された第一の閾値との比較を行い、歪の方が大きい場合には比較結果として“1”を、それ以外の場合には比較結果として“0”を出力する。補正手段 3 1 は、判定手段 1 4 から出力された判定結果と、比較手段 3 0 から出力された比較結果の両方が“1”である場合に、駆動音源符号化手段 1 0 から出力された結果中の歪を閾値計算手段 2 9 から入力された第一の閾値を用いて補正し、補正後の値を新たに歪として最小歪選択手段 1 7 に出力する。なお、それ以外の場合には上記補正は行わず、駆動音源符号化手段 1 0 から出力された歪をそのまま最小歪選択手段 1 7 に出力する。この補正手段 3 1 による補正については、歪を  $D$ 、閾値を  $D_{th}$  とすれば、例えば次の (6) 式を用いて行うことができる。

$$D' = D + \alpha (D - D_{th}) \quad \dots (6)$$

ここで、 $D'$  は補正後の歪、 $\alpha$  は正の定数である。

## 【0 1 3 1】

なお、補正手段 3 1 による補正は、指数関数を適用するなど、(6) 式よりも複雑な補正を適用することも当然可能であるし、非常に大きい固定値に補正してしまうことも可能である。非常に大きい固定値に補正した場合には、最小歪選択手段 1 7 において基本的に駆動音源符号化手段 1 0 が選択されなくなる。

## 【0 1 3 2】

また、比較手段 3 2 は、駆動音源符号化手段 1 1 から入力された歪と、閾値計算手段 2 9 から入力された第二の閾値を比較し、歪の方が大きい場合には比較結果として“1”を、それ以外の場合には比較結果として“0”を出力する。補正手段 3 3 は、判定手段 1 4 から出力された判定結果と、比較手段 3 2 から出力された比較結果の両方が“1”である場合に、駆動音源符号化手段 1 1 から出力された結果中の歪を、閾値計算手段 2 9 から入力された閾値を用いて補正し、補正後の値を新たに歪として最小歪選択手段 1 7 に出力する。それ以外の場合には補正を行わず、駆動音源符号化手段 1 0 からの歪をそのまま最小歪選択手段 1 7 に出力する。なお、この補正については、補正手段 3 1 の場合と同様に行うことができる。

#### 【0 1 3 3】

最小歪選択手段 1 7 は、駆動音源符号化手段 9、補正手段 3 1、および補正手段 3 3 から入力された各歪の比較を行い、それらの中の最小の歪を選択する。その結果、駆動音源符号化手段 9 から入力された歪を選択した場合には、駆動音源符号化手段 9 から入力された駆動音源をゲイン符号化手段 6 に、駆動音源符号を多重化手段 7 にそれぞれ出力する。また補正手段 3 1 から入力された歪を選択した場合には、補正手段 3 1 を介して駆動音源符号化手段 1 0 から入力された駆動音源をゲイン符号化手段 6 に、駆動音源符号を多重化手段 7 にそれぞれ出力する。同様に、補正手段 3 3 から入力された歪を選択した場合には、補正手段 3 3 を介して駆動音源符号化手段 1 1 から入力された駆動音源をゲイン符号化手段 6 に、駆動音源符号を多重化手段 7 にそれぞれ出力する。さらに、これら 3 つの歪の内のどれを選択したかを示す情報を、モード選択情報として多重化手段 7 に出力する。

#### 【0 1 3 4】

次に、この実施の形態 3 により得られた音声符号 8 を音声復号化装置で復号化して得られる復号音の主観的な品質（音質）が改善されることを、図 7 を参照しながら説明する。

図 7 は符号化歪を最小にする音源モードの選択について説明するための各波形を示すイメージ図で、図 7（a）は入力音声、同図（b）は雑音的な音声を表現

するために用意した音源モードを選択したときの復号音、同図（c）は母音的な音声を表現するために用意した音源モードを選択したときの復号音をそれぞれ示している。入力音声 1 が図 7（a）に示すように雑音的である場合には、一般にモデル化がうまく働かないので、同図（b）に示した雑音的な音声を表現するために用意した音源モードの場合でも、同図（c）に示した母音的な音声を表現するために用意した音源モードの場合でも、符号化時の歪率は比較的大きな値となる。

## 【0 1 3 5】

ここで、駆動音源符号化手段 9 は乱数によって生成された時系列ベクトルを使用しており、図 7（b）に示す雑音的な音声を表現するために用意した音源モードに対応する。また駆動音源符号化手段 1 0 および 1 1 はパルス音源とピッチ周期化を使用しており、図 7（c）に示す母音的な音声を表現するために用意した音源モードに対応する。

## 【0 1 3 6】

各駆動音源符号化手段 9 ～ 1 1 より出力された歪  $D$  は何れも大きな値となっているが、駆動音源符号化手段 1 0 および 1 1 から出力された歪  $D$  は、補正手段 3 1 あるいは 3 3 によって、歪  $D$  よりも大きい値に補正される。その結果、駆動音源符号化手段 9 が出力した駆動音源符号が、最小歪選択手段 1 7 によって選択され、復号音は図 7（b）に示すものとなる。このように、図 7（b）の歪の方が同図（c）の歪より大きくても、雑音的な区間などの符号化時の歪率が大きくなる区間では、安定的に図 7（b）に示す復号音が選択される。

## 【0 1 3 7】

なお、この実施の形態 3 では、各駆動音源符号化手段 9 ～ 1 1 が、（1）式に示した歪  $D$  を最小化する駆動音源符号を探索し、最小の歪  $D$  を出力するようにするものについて説明したが、実施の形態 1 の場合と同様に、（3）式に示した評価値  $d$  を最大化する駆動音源符号を探索し、歪  $D$  の代わりに評価値  $d$  を出力する構成とすることも可能である。

## 【0 1 3 8】

また、この実施の形態 3 では、閾値計算手段 2 9 が 2 つの固定閾値をそのまま

出力し、各駆動音源符号化手段 9～11 が出力する歪を入力音声 1 の信号パワーで除算した値、つまり歪率として出力するように変形することも可能であるし、パワー計算手段 12 が適応音源符号化手段 4 の出力する符号化対象信号の信号パワーを計算するように変更することも、信号パワーの代わりに、振幅や、対数パワーなどを計算するように変更することも可能である。

## 【0139】

また、この実施の形態 3 では、雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段 9 を 1 つ、また非雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段 10、11 の 2 つを備える構成としたが、前者を 2 つ以上としてもかまわないし、後者を 1 つまたは 3 つ以上としてもかまわないことはいうまでもない。

## 【0140】

また、この実施の形態 3 では、単純に信号間の 2 乗距離を歪としているが、音声符号化装置で多く使用されている聴覚重み付け歪とするようにしても当然かまわない。

## 【0141】

以上のように、この実施の形態 3 によれば、実施の形態 1 の場合と同様に、符号化歪が大きい場合、あるいは符号化時の歪率が所定以上である場合でも、復号音の品質劣化が少ない音源モードを選択することが可能となり、また符号化歪が大きくても復号音の品質劣化を起こしにくい入力音声に対しては、従来の場合と同様の音源モード選択が行われて一層丁寧な音源モード選択が可能となり、また符号化歪が大きくなりがちな区間とそれ以外の区間で符号化歪に基づく音源モード選択の制御を変えることができ、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することが可能となり、さらに符号化歪が大きい場合に、雑音的な音源を生成する音源モード、もしくは雑音的な音源符号語を使用する音源モードが選択しやすくなって、非雑音的な音源を生成する音源モード、あるいは非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避することが可能となるなどのため、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化した復号音の主観的な品質（音質）を改



善することができるという効果が得られる。

【0142】

さらに、この実施の形態3によれば、符号化歪が閾値を上回る比較結果が得られている音源モードの選択を抑制するようにしたので、符号化歪が大きい場合に復号音の品質劣化が少ない音源モードを選択しやすくすることが可能となり、よりよい音質を与える音源モードを適切に選択できるため、得られた音声符号を復号化した復号音の主観的な品質（音質）を改善することができるという効果が得られる。

【0143】

さらに、この実施の形態3によれば、音源モード毎に閾値を用意したので、各音源モード毎に復号音品質の劣化を引き起こすことを検知する閾値を適切に調整することで、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化して得られる復号音の主観的な品質（音質）を改善することができるという効果が得られる。

【0144】

実施の形態4.

図4はこの発明の実施の形態4による音声符号化方法を適用した音声符号化装置の構成を示すブロック図であり、図中、図1の各部分と同等の部分については同一番号を付してその説明を省略する。図において、34は入力音声1と、線形予測係数符号化手段3からの信号および適応音源符号化手段4からの信号をもとに、駆動音源と駆動音源符号、およびモード選択情報を出力する駆動音源符号化部である。

【0145】

また、35は駆動音源符号化手段9～11からの信号に基づいて最小の歪と、この最小の歪に対応する駆動音源と、駆動音源符号と、モード選択情報を出力する、選択手段としての最小歪選択手段である。36はこの最小歪選択手段35からの信号と閾値計算手段13からの閾値とを比較する比較手段であり、37はこの比較手段36と判定手段14の判定結果に基づいて最小歪選択手段35からの信号を駆動音源符号化手段9の出力で置換する置換手段である。なお、上記駆動

音源符号化部 34 は、最小歪選択手段 35、比較手段 36、置換手段 37 と、駆動音源符号化手段 9、10、11、パワー計算手段 12、閾値計算手段 13、および判定手段 14 とによって構成されている。

【0146】

次に動作について説明する。なお、ここでは、上記実施の形態 1 とは異なる部分を中心に、図 4 に基づいて説明する。

この場合も、線形予測係数符号化手段 3 で量子化された線形予測係数と、適応音源符号化手段 4 からの符号化対象信号とが、駆動音源符号化部 34 内の駆動音源符号化手段 9～11 に入力されている。この駆動音源符号化手段 9 には、乱数によって生成された複数の時系列ベクトルが駆動音源符号帳として記憶されている。駆動音源符号化手段 9 は実施の形態 1 の場合と同様に、その駆動音源符号帳を用いて適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、最小歪選択手段 35 と置換手段 37 とに出力する。

【0147】

また駆動音源符号化手段 10 には、パルス位置テーブルを含む駆動音源符号帳が記憶されている。駆動音源符号化手段 10 はその駆動音源符号帳を用いて、適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、最小歪選択手段 35 に出力する。同様に、駆動音源符号化手段 11 には、上記駆動音源符号化手段 10 のものとは異なるパルス位置テーブルを含む駆動音源符号帳が記憶されている。駆動音源符号化手段 10、11 はそれらの駆動音源符号帳を用いて、適応音源符号化手段 4 から入力された符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、最小歪選択手段 35 に出力する。

【0148】

なお、この場合も、駆動音源符号化手段 9 の駆動音源符号帳には、乱数によって生成された雑音的な音源符号語が、駆動音源符号化手段 1 0, 1 1 の駆動音源符号帳には、パルス位置テーブル等による非雑音的な音源符号語がそれぞれ格納されており、また、駆動音源符号化手段 9 より出力される時系列ベクトルは、雑音的な音源を生成し、駆動音源符号化手段 1 0, 1 1 より出力される時系列ベクトルは、非雑音的な音源を生成している。

## 【 0 1 4 9 】

最小歪選択手段 3 5 はこれら各駆動音源符号化手段 9 ～ 1 1 より入力された各歪の比較を行い、それらの中の最小の歪を選択して、その最小の歪を比較手段 3 6 に出力する。また、駆動音源符号化手段 9 ～ 1 1 中の、この最小の歪に対応するものから入力された駆動音源と駆動音源符号を置換手段 3 7 に出力し、さらにこれら 3 つの歪の内のどれを選択したかを示す情報を、モード選択情報として置換手段 3 7 に出力する。また、判定手段 1 4 は入力音声 1 を分析して音声態様の判定を行い、音声の立ち上がりの場合には“0”を、それ以外の場合には“1”を、判定結果として置換手段 3 7 に出力する。

## 【 0 1 5 0 】

一方、比較手段 3 6 には上記最小歪選択手段 3 5 で選択された歪とともに、閾値計算手段 1 3 にてパワー計算手段 1 2 からの信号パワーに基づいて計された歪に関する閾値が入力されている。また、比較手段 3 6 は最小歪選択手段 3 5 から入力された歪と、閾値計算手段 1 3 から入力された閾値とを比較して、歪の方が大きい場合には“1”を、それ以外の場合には“0”を、比較結果として置換手段 3 7 に出力する。

## 【 0 1 5 1 】

置換手段 3 7 は、この判定手段 1 4 から出力された判定結果と比較手段 3 6 から出力された比較結果とを受け、その両方が“1”であった場合には、最小歪選択手段 3 5 から出力された駆動音源と駆動音源符号を、駆動音源符号化手段 9 から出力された駆動音源と駆動音源符号に置換する。なお、それ以外の場合には上記置換を行わない。この置換手段 3 7 による置換処理結果である最終的な駆動音源がゲイン符号化手段 6 に、駆動音源符号が多重化手段 7 に出力される。

## 【0152】

次に、この実施の形態4により得られた音声符号8を音声復号化装置で復号化して得られる復号音の主観的な品質（音質）が改善されることを、図7を参照しながら説明する。

図7は符号化歪を最小にする音源モードの選択について説明するための各波形を示すイメージ図で、図7（a）は入力音声、同図（b）は雑音的な音声を表現するために用意した音源モードを選択したときの復号音、同図（c）は母音的な音声を表現するために用意した音源モードを選択したときの復号音をそれぞれ示している。入力音声1が図7（a）に示すように雑音的である場合には、一般にモデル化がうまく働かないので、同図（b）に示した雑音的な音声を表現するために用意した音源モードの場合でも、同図（c）に示した母音的な音声を表現するために用意した音源モードの場合でも、符号化時の歪率は比較的大きな値となる。

## 【0153】

ここで、駆動音源符号化手段9は乱数によって生成された時系列ベクトルを使用しており、図7（b）に示す雑音的な音声を表現するために用意した音源モードに対応する。また駆動音源符号化手段10および11はパルス音源とピッチ周期化を使用しており、図7（c）に示す母音的な音声を表現するために用意した音源モードに対応する。

## 【0154】

各駆動音源符号化手段9～11より出力された歪Dは何れも大きな値となっているが、最小歪選択手段35では、振幅の大きい部分での符号化歪が少ないために歪Dが小さくなることの多い、駆動音源符号化手段10または11から出力された歪が選択される。それでも選択された最小の歪Dの値は閾値計算手段13からの閾値 $D_{th}$ よりは大きいため、最小歪選択手段35より出力された駆動音源符号化手段10または11の駆動音源符号は、置換手段37にて駆動音源符号化手段9より出力された駆動音源符号に置換されて、復号音は図7（b）に示すものとなる。このように、図7（b）の歪の方が同図（c）の歪より大きくても、雑音的な区間などの符号化時の歪率が大きくなる区間では、図7（b）に示す復号

音が安定的に選択される。

【0155】

なお、この実施の形態4においても、実施の形態1と同様に、各駆動音源符号化手段9～11が、(3)式に示した評価値 $d$ を最大化する駆動音源符号を探索し、歪 $D$ の代わりに評価値 $d$ を出力する構成とすることも可能である。この場合、最小歪選択手段35では最大の評価値を選択し、比較手段36では大小関係と出力結果の関係が逆になる。また閾値計算手段13でも評価値 $d$ に対応した閾値 $d_{th}$ を計算する必要がある。

【0156】

また、この実施の形態4では、閾値計算手段13が固定の閾値をそのまま出力し、各駆動音源符号化手段9～11が出力する歪を入力音声1の信号パワーで除算した値、つまり歪率として出力するように変形することも可能であるし、パワー計算手段12が適応音源符号化手段4の出力する符号化対象信号の信号パワーを計算するように変更することも、信号パワーの代わりに、振幅や、対数パワーなどを計算するように変更することも可能である。

【0157】

また、この実施の形態4では、雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段9を1つ、また非雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段10、11の2つを備える構成としたが、前者を2つ以上としてもかまわないし、後者を1つまたは3つ以上としてもかまわないことはいうまでもない。

【0158】

また、この実施の形態4では、単純に信号間の2乗距離を歪としているが、音声符号化装置で多く使用されている聴覚重み付け歪とするようにしても当然かまわない。

【0159】

以上のように、この実施の形態4によれば、複数の音源モードの中から1つを選択し、この音源モードを使用して入力音声1をフレームと呼ばれる所定長の区間毎に符号化する際に、各音源モード毎に入力音声から求まる符号化対象信号の

符号化を行ってそれらの中の1つを選択し、それと、その時の符号化歪と固定の閾値、あるいは符号化対象信号の信号パワーに基づいて決定した閾値との比較を行い、その比較結果に基づいて符号化歪の出力変換を行うようにしたので、符号化歪が大きい場合でも復号音の品質劣化が少ない音源モードを選択することが可能となるため、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質を改善することができるという効果が得られる。

## 【0160】

さらに、この実施の形態4によれば、実施の形態1の場合と同様に、符号化時の歪率が所定以上である場合でも、復号音の品質劣化が少ない音源モードを選択することが可能となり、また符号化歪が大きくても復号音の品質劣化を起こしにくい入力音声に対しては、従来の場合と同様の音源モード選択が行われて一層丁寧な音源モード選択が可能となり、また符号化歪が大きくなりがちな区間とそれ以外の区間で符号化歪に基づく音源モード選択の制御を変えることができ、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することが可能となり、さらに符号化歪が大きい場合に、雑音的な音源を生成する音源モード、もしくは雑音的な音源符号語を使用する音源モードが選択しやすくなって、非雑音的な音源を生成する音源モード、あるいは非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避することが可能となるなどのため、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化した復号音の主観的な品質を改善することができるという効果が得られる。

## 【0161】

さらに、この実施の形態4によれば、符号化歪の中で最小のものを選択し、選択した符号化歪と閾値の比較を行い、その比較結果に基づいて、音源モードの選択を行うようにしたので、符号化歪が大きい場合に復号音の品質劣化が少ない音源モードを強制的に選択することが可能となって、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化して得られる復号音の主観的な品質を改善することができるという効果が得られる。

## 【0162】

さらに、この実施の形態 4 によれば、符号化歪の中で最小のものを選択し、選択した符号化歪が閾値を上回る場合には、あらかじめ定めた音源モードを選択するようにしたので、符号化歪が大きい場合に復号音の品質劣化が少ない音源モードを強制的に選択することが可能となって、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化して得られる復号音の主観的な品質を改善することができる効果が得られる。

## 【 0 1 6 3 】

実施の形態 5.

図 5 はこの発明の実施の形態 5 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図であり、図中、図 1 の各部分と同等の部分については同一番号を付してその説明を省略する。図において、38 は入力音声 1 と、線形予測係数符号化手段 3 からの信号および適応音源符号化手段 4 からの信号をもとに、駆動音源と駆動音源符号、およびモード選択情報を出力する駆動音源符号化部である。

## 【 0 1 6 4 】

また、39 は入力音声 1 を分析し、音声の立ち上がり部分であるか否かの判定を行う判定手段であるが、判定結果を変換手段ではなく閾値計算手段に出力している点で、図 1 に符号 14 を付したものと異なっている。40 はこの判定手段 39 の判定結果とパワー計算手段 12 からの信号パワーをもとに閾値を算出する閾値計算手段である。41 は判定手段 39 の判定結果と比較手段 15 の比較結果に基づいて駆動音源符号化手段 9 の出力変換を行う変換手段である。なお、上記駆動音源符号化部 38 は、判定手段 39、閾値計算手段 40、変換手段 41 と、駆動音源符号化手段 9 ~ 11、パワー計算手段 12、比較手段 15、および最小歪選択手段 17 とによって構成されている。

## 【 0 1 6 5 】

次に動作について説明する。なお、ここでは、上記実施の形態 1 とは異なる部分を中心に、図 5 に基づいて説明する。

この場合も、線形予測係数符号化手段 3 で量子化された線形予測係数と、適応音源符号化手段 4 からの符号化対象信号とが、駆動音源符号化部 38 内の駆動音

源符号化手段 9 ～ 11 に入力されており、駆動音源符号化手段 9 は、乱数によって生成された複数の時系列ベクトルが格納された駆動音源符号帳を用いて、符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、変換手段 41 と比較手段 15 とに出力する。また、駆動音源符号化手段 10 および 11 は、互いに異なったパルス位置テーブルを含む駆動音源符号帳を用いて、符号化対象信号を符号化した時の歪を最小にする駆動音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、最小の歪および駆動音源符号とともに、最小歪選択手段 17 に出力する。

## 【0166】

なお、この場合も、駆動音源符号化手段 9 の駆動音源符号帳には、乱数によって生成された雑音的な音源符号語が、駆動音源符号化手段 10, 11 の駆動音源符号帳には、パルス位置テーブル等による非雑音的な音源符号語がそれぞれ格納されており、また、駆動音源符号化手段 9 より出力される時系列ベクトルは、雑音的な音源を生成し、駆動音源符号化手段 10, 11 より出力される時系列ベクトルは、非雑音的な音源を生成している。

## 【0167】

一方、パワー計算手段 12 は入力音声 1 のフレーム内の信号パワーを計算して閾値計算手段 40 に出力する。また、判定手段 39 は入力音声 1 を分析して音声態様の判定を行い、判定の結果、音声の立ち上がり部分であった場合には“0”を、それ以外の場合には“1”を、それぞれ閾値計算手段 40 に出力する。

## 【0168】

閾値計算手段 40 は、判定手段 39 の判定結果が“0”であった場合には、あらかじめ用意されていた歪率に関する第一の定数をパワー計算手段 12 からの信号パワーに乗算し、判定手段 39 の判定結果が“1”であった場合には、あらかじめ用意した歪率に関する第二の定数をパワー計算手段 12 から入力された信号パワーに乗算する。この乗算によって得られた結果を歪に関する閾値として、それぞれ比較手段 15 と変換手段 41 に出力する。なお、上記第一の定数は第二の定数に比べて大きく設定しておく。例えば、第一の定数を 0.9、第二の定数を



0. 7に設定する。

【0169】

比較手段15は、駆動音源符号化手段9から入力された歪と、閾値計算手段40から入力された閾値との比較を行い、歪の方が大きい場合時には“1”を、それ以外の場合には“0”を比較結果として変換手段41に出力する。変換手段41は、この比較手段15から出力された比較結果が“1”である場合には、駆動音源符号化手段9から出力された結果中の歪を、閾値計算手段40から入力された閾値の値に置換して最小歪選択手段17に出力する。また、それ以外の場合には、上記置換を行わず、駆動音源符号化手段9から出力された結果中の歪をそのまま、最小歪選択手段17に出力する。

【0170】

最小歪選択手段17は、変換手段41から入力された歪と、駆動音源符号化手段10および11から入力された歪との比較を行って、それらの中で最小の歪を選択する。そしてこの選択された最小の歪を出力した、変換手段41または駆動音源符号化手段10または駆動音源符号化手段11からの駆動音源をゲイン符号化手段6に、駆動音源符号を多重化手段7にそれぞれ出力する。さらに、これら3つの歪の内のどれを選択したかを示す情報を、モード選択情報として多重化手段7に出力する。

【0171】

次に、この実施の形態3により得られた音声符号8を音声復号化装置で復号化して得られる復号音の主観的な品質（音質）が最適の選択となることを、図7を参照しながら説明する。

図7は符号化歪を最小にする音源モードの選択について説明するための各波形を示すイメージ図であり、入力音声1が図7（a）に示すように雑音的である場合には、一般にモデル化がうまく働かないので、同図（b）に示した雑音的な音声を表現するために用意した音源モードの場合でも、同図（c）に示した母音的な音声を表現するために用意した音源モードの場合でも、符号化時の歪率は比較的大きな値となる。

【0172】

ここで、駆動音源符号化手段 9 は乱数によって生成された時系列ベクトルを使用しており、図 7 (b) に示す雑音的な音声を表現するために用意した音源モードに対応する。また駆動音源符号化手段 1 0 および 1 1 はパルス音源とピッチ周期化を使用しており、図 7 (c) に示す母音的な音声を表現するために用意した音源モードに対応する。

## 【 0 1 7 3 】

判定手段 3 9 が音声の立ち上がりであると判定して、判定結果 “0” を出力した場合には、閾値計算手段 4 0 では比較的大きな閾値を算出する。このため、駆動音源符号化手段 9 から出力された歪  $D$  は大きい値ではあるが、それでも閾値を上回るには至らず、変換手段 4 1 での置換処理は行われぬ。その結果、最小歪選択手段 1 7 において、振幅の大きい部分での符号化歪が少ないために歪  $D$  が小さくなることが多い駆動音源符号化手段 1 0 または 1 1 が選択されて、復号音は図 7 (c) に示すものとなる。

## 【 0 1 7 4 】

判定手段 3 9 が音声の立ち上がり以外であると判定して、判定結果 “1” を出力した場合には、閾値計算手段 4 0 では比較的小さな閾値を算出する。このため、駆動音源符号化手段 9 から出力された歪  $D$  が閾値を上回り、変換手段 4 1 での歪  $D$  よりも小さい閾値  $D_{th}$  に置換される。その結果、最小歪選択手段 1 7 において、駆動音源符号化手段 9 が出力した駆動音源符号が選択されて、復号音は図 7 (b) に示すものとなる。このように、図 7 (b) の歪の方が同図 (c) の歪より大きくても、音声の立ち上がり以外で、雑音的な区間など符号化時の歪率が大きくなる区間では、図 7 (b) に示す復号音が安定的に選択される。

## 【 0 1 7 5 】

なお、音声の立ち上がりでも比較的小さな閾値を用いて復号音が図 7 (b) に示すものとなるようにしてしまうと、破裂音のパルスの特徴がつぶれてしまったり、母音の立ち上がりをザラザラした音質に劣化させてしまうことになる。この実施の形態 5 では判定手段 3 9 での判定結果に基づいて閾値を決定することにより、立ち上がりでの劣化を回避している。

## 【 0 1 7 6 】

なお、この実施の形態5についても、実施の形態1と同様に、各駆動音源符号化手段9～11が、(3)式に示した評価値 $d$ を最大化する駆動音源符号を探索し、歪 $D$ の代わりに評価値 $d$ を出力する構成とすることも可能である。この場合、最小歪選択手段17では最大の評価値を選択し、比較手段15では大小関係と出力結果の関係が逆になる。また閾値計算手段40でも評価値 $d$ に対応した閾値 $d_{th}$ を計算する必要がある。

## 【0177】

また、この実施の形態5では、閾値計算手段40が第一の定数または第二の定数をそのまま閾値として出力し、各駆動音源符号化手段9～11が出力する歪を入力音声1の信号パワーで除算した値、つまり歪率として出力するように変形することも可能であるし、パワー計算手段12が適応音源符号化手段4の出力する符号化対象信号の信号パワーを計算するように変更することも、信号パワーの代わりに、振幅や、対数パワーなどを計算するように変更することも可能である。

## 【0178】

また、この実施の形態5では、雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段9を1つ、また非雑音的な音源を生成する駆動音源符号化手段として駆動音源符号化手段10、11の2つを備える構成としたが、前者を2つ以上としてもかまわないし、後者を1つまたは3つ以上としてもかまわないことはいうまでもない。

## 【0179】

また、この実施の形態5では、単純に信号間の2乗距離を歪としているが、音声符号化装置で多く使用されている聴覚重み付け歪とするようにしても当然かまわない。

## 【0180】

また、この実施の形態5では、閾値計算手段40が判定手段39の判定結果に基づいて、あらかじめ用意した歪率に関する2つの定数の一方を選択して使用するようにしているが、判定結果を3つ以上とすることで、対応する定数も3つ以上とし、より細かい制御を行うようにすることも可能である。さらに、判定手段39が入力音声1を分析して連続値の判定パラメータを算出し、閾値計算手段4

0がこの判定パラメータに基づいて、連続値による閾値を計算するようにすることも可能である。

#### 【0181】

以上のように、この実施の形態5によれば、実施の形態1の場合と同様に、符号化歪が大きい場合、あるいは符号化時の歪率が所定以上である場合でも、復号音の品質劣化が少ない音源モードを選択することが可能となり、また符号化歪が大きい場合でも、符号化歪が置換された音源モードが選択されやすくなり、また符号化歪が大きくなりがちな区間とそれ以外の区間で符号化歪に基づく音源モード選択の制御を変えることができ、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することが可能となり、さらに符号化歪が大きい場合に、雑音的な音源を生成する音源モード、もしくは雑音的な音源符号語を使用する音源モードが選択しやすくなって、非雑音的な音源を生成する音源モード、あるいは非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避することが可能となるなどのため、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化した復号音の主観的な品質を改善することができるという効果が得られる。

#### 【0182】

さらに、この実施の形態5によれば、入力音声1または符号化対象信号の分析を行って音声様態を判定し、その判定結果に基づいて決定した閾値を用いて比較を行うようにしたので、音声様態に応じて適切に設定された閾値を用いて音源モードの選択を行うことが可能となるため、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質を改善することができるという効果が得られる。

#### 【0183】

実施の形態6.

図6はこの発明の実施の形態6による音声符号化方法を適用した音声符号化装置の構成を示すブロック図であり、図中、図1の各部分と同等の部分については同一番号を付してその説明を省略する。図において、42は入力音声1と、線形予測係数符号化手段3からの信号および適応音源符号化手段4からの信号をもと

に、駆動音源と駆動音源符号、およびモード選択情報を出力する駆動音源符号化部である。

#### 【0184】

また、43は乱数によって生成された時系列ベクトルによる駆動音源符号帳であり、44はこの駆動音源符号帳43を用いて、線形予測係数符号化手段3と適応音源符号化手段4からの信号より仮の合成音と符号化対象信号との歪を調べ、駆動音源を出力する符号化手段としての駆動音源符号化手段である。45はパルス位置テーブルを含む駆動音源符号帳であり、46はこの駆動音源符号帳45を用いて、線形予測係数符号化手段3と適応音源符号化手段4からの信号より仮の合成音と符号化対象信号との歪を調べ、駆動音源を出力する符号化手段としての駆動音源符号化手段である。なお、上記駆動音源符号化部42は、パワー計算手段12、閾値計算手段13、判定手段14、比較手段15、変換手段16、最小歪選択手段17と、駆動音源符号帳43、45、駆動音源符号化手段44、46とによって構成されている。

#### 【0185】

次に動作について説明する。なお、ここでは、上記実施の形態1とは異なる部分を中心に、図6に基づいて説明する。

駆動音源符号帳43には乱数によって生成された複数の時系列ベクトルが記憶されている。この駆動音源符号帳43は数ビットの2進数値で示した駆動音源符号が入力されると、その駆動音源符号に対応する位置に格納してある時系列ベクトルを読み出して出力する。駆動音源符号化手段44はこの駆動音源符号帳43に各駆動音源符号を入力することによって得られる時系列ベクトルを、線形予測係数符号化手段3の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と、適応音源符号化手段4から入力された符号化対象信号との差分をとり、それに基づく両者間の歪を調べる。

#### 【0186】

駆動音源符号化手段44はこの処理を全ての音源符号に対して行い、最小の歪を与えた音源符号を選択し、その選択された音源符号に対応する時系列ベクトル

を駆動音源として、上記最小の歪および音源符号とともに、比較手段 1 5 と変換手段 1 6 とに出力する。

## 【 0 1 8 7 】

また、駆動音源符号帳 4 5 にはパルス位置テーブルを含む符号帳が記憶されている。この駆動音源符号帳 4 5 は数ビットの 2 進数値で示した駆動音源符号が入力されると、その駆動音源符号を複数のパルス位置符号と極性に分離し、パルス位置テーブル中の各パルス位置符号に対応する位置に格納してあるパルス位置を読み出して、このパルス位置と極性に従って複数本のパルスを持った時系列ベクトルを生成する。駆動音源符号帳 4 5 はさらに、その生成された時系列ベクトルを、適応音源符号化手段 4 で選択された適応音源符号に対応する繰返し周期を用いてピッチ周期化し、駆動音源符号化手段 4 6 に出力する。

## 【 0 1 8 8 】

駆動音源符号化手段 4 6 は駆動音源符号帳 4 5 に各駆動音源符号を入力することによって得られる各時系列ベクトルを、線形予測係数符号化手段 3 の出力する、量子化された線形予測係数を用いた合成フィルタにてフィルタリングすることにより、仮の合成音を求める。そして、得られた仮の合成音に適切なゲインを乗算した信号と、適応音源符号化手段 4 から入力された符号化対象信号との差分をとり、それに基づく両者間の歪を調べる。駆動音源符号化手段 4 6 はこの処理を全ての音源符号に対して行い、最小の歪を与えた音源符号を選択し、その選択された音源符号に対応する時系列ベクトルを駆動音源として、上記最小の歪および音源符号とともに、最小歪選択手段 1 7 に出力する。

## 【 0 1 8 9 】

なお、この場合も、駆動音源符号化手段 4 4 の駆動音源符号帳 4 3 には、乱数によって生成された雑音的な音源符号語が、駆動音源符号化手段 4 6 の駆動音源符号帳 4 5 には、パルス位置テーブル等による非雑音的な音源符号語がそれぞれ格納されており、また、駆動音源符号化手段 4 4 より出力される時系列ベクトルは、雑音的な音源を生成し、駆動音源符号化手段 4 6 より出力される時系列ベクトルは、非雑音的な音源を生成している。

## 【 0 1 9 0 】

一方、パワー計算手段 1 2 は入力音声 1 のフレーム内の信号パワーを計算して閾値計算手段 1 3 に出力する。閾値計算手段 1 3 は、あらかじめ用意されている歪率に関する定数を、このパワー計算手段 1 2 から入力された信号パワーに乗算し、得られた結果を歪に関する閾値として比較手段 1 5 および変換手段 1 6 に出力する。また、判定手段 1 4 は入力音声 1 を分析して音声態様の判定を行い、判定の結果、音声の立ち上がり部分であった場合には“0”を、それ以外の場合には“1”を、それぞれ閾値計算手段 1 3 に出力する。

## 【0 1 9 1】

比較手段 1 5 は、駆動音源符号化手段 4 4 から入力された歪と、閾値計算手段 1 3 から入力された閾値を比較する。比較の結果、歪の方が大きかった場合には“1”を、それ以外の場合には“0”を、比較結果として変換手段 1 6 に出力する。変換手段 1 6 はこの判定手段 1 4 から出力された判定結果と、比較手段 1 5 から出力された比較結果の両方が“1”であれば、駆動音源符号化手段 4 4 から入力された結果中の歪を、閾値計算手段 1 3 から入力された閾値の値に置換して最小歪選択手段 1 7 に出力する。なお、それ以外の場合には上記置換は行わず、駆動音源符号化手段 4 4 から入力された歪をそのまま最小歪選択手段 1 7 に出力する。

## 【0 1 9 2】

最小歪選択手段 1 7 はこの変換手段 1 6 から入力された歪と、駆動音源符号化手段 4 6 から入力された歪とを比較して、その小さい方の歪を選択する。そしてこの選択された最小の歪を出力した変換手段 1 6 または駆動音源符号化手段 4 6 からの駆動音源をゲイン符号化手段 6 に、駆動音源符号を多重化手段 7 にそれぞれ出力する。さらに、これら 2 つの歪のどちらを選択したかを示す情報を、モード選択情報として多重化手段 7 に出力する。

## 【0 1 9 3】

ここで、この駆動音源符号化手段 4 4 と駆動音源符号化手段 4 6 における符号化処理の差異は、アクセスする駆動音源符号帳 4 3、4 5 の違いのみである。このような場合、駆動音源符号帳 4 3 と駆動音源符号帳 4 5 を一体化して、1 つの駆動音源符号化手段によって探索するように変形することができる。そのような

場合には、駆動音源符号帳 43 に対応する駆動音源による歪と、駆動音源符号帳 45 に対応する歪を独立に計算し、前者の歪を変換手段 16 に入力することで同じ結果を得ることができる。つまり、1つの駆動音源符号帳に対する駆動音源符号を、雑音的な符号語に対応するものと、非雑音的な符号語に対応するものとにグループ分けし、前者を駆動音源符号帳 43、後者を駆動音源符号帳 45 と見なせば、この実施の形態 6 を適用することが可能となる。

## 【0194】

なお、この実施の形態 6 についても実施の形態 1 と同様に、駆動音源符号化手段 44、46 が (3) 式に示す評価値  $d$  を最大化する駆動音源符号を探索し、歪  $D$  の代わりに評価値  $d$  を出力する構成とすることも可能である。その場合、最小歪選択手段 17 では最大の評価値を選択し、比較手段 15 では大小関係と出力結果の関係が逆になる。また閾値計算手段 13 でも評価値  $d$  に対応した閾値  $d_{th}$  を計算する必要がある。

## 【0195】

また、この実施の形態 6 では、閾値計算手段 13 が歪率に関する定数をそのまま閾値として出力し、各駆動音源符号化手段 44、46 が出力する歪を入力音声 1 の信号パワーで除算した値、つまり歪率として出力するように変形することも可能であり、パワー計算手段 12 が適応音源符号化手段 4 の出力する符号化対象信号の信号パワーを計算するように変更することも、信号パワーの代わりに、振幅や、対数パワーなどを計算するように変更することも可能である。

## 【0196】

また、この実施の形態 6 では、雑音的な音源を生成する駆動音源符号化手段としての駆動音源符号化手段 44 と、非雑音的な音源を生成する駆動音源符号化手段としての駆動音源符号化手段 46 をそれぞれ 1 つずつ備える構成としたが、それぞれを 2 つ以上としてもかまわないことはいうまでもない。

## 【0197】

また、この実施の形態 6 では、単純に信号間の 2 乗距離を歪としているが、音声符号化装置で多く使用されている聴覚重み付け歪とするようにしても当然かまわない。



## 【0198】

以上のように、この実施の形態6によれば、実施の形態1の場合と同様に、符号化歪が大きい場合、あるいは符号化時の歪率が所定以上である場合でも、復号音の品質劣化が少ない音源モードを選択することが可能となり、また符号化歪が大きい場合でも、符号化歪が置換された音源モードが選択されやすくなり、また符号化歪が大きくても復号音の品質劣化を起こしにくい入力音声に対しては、従来の場合と同様の音源モード選択が行われて一層丁寧な音源モード選択が可能となり、また符号化歪が大きくなりがちな区間とそれ以外の区間で符号化歪に基づく音源モード選択の制御を変えることができ、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することが可能となり、さらに符号化歪が大きい場合に、雑音的な音源を生成する音源モード、もしくは雑音的な音源符号語を使用する音源モードが選択しやすくなって、非雑音的な音源を生成する音源モード、あるいは非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避することが可能となるなどのため、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化した復号音の主観的な品質を改善することができるという効果が得られる。

## 【0199】

実施の形態7.

なお、上記実施の形態2では、適応音源符号化手段と駆動音源符号化手段によって構成される複数の音源符号化手段19～21を備えて、その中の1つを選択するものについて説明したが、ゲイン符号化手段6も含むさらに上位の音源符号化手段を複数備えて、その中の1つを選択するようにしてもよい。

## 【0200】

また、実施の形態3～実施の形態6についても、適応音源符号化手段4と駆動音源符号化手段9～11あるいは44、46によって構成される複数の音源符号化手段を備えて、その中の1つを選択するように構成しても、ゲイン符号化手段6も含むさらに上位の音源符号化手段を複数備えて、その中の1つを選択するように構成するようにしてもよい。

## 【0201】

このように上位の音源モードを複数備えて、この音源モードを使用して入力音声をフレームと呼ばれる所定長区間毎に符号化する音声符号化方法において、各音源モード毎に入力音声から求まる符号化対象信号の符号化を行い、その時の符号化歪と固定の閾値または符号化対象信号の信号パワーに基づいて決定した閾値との比較を行い、その比較結果に基づいて、音源モードの選択を行うようにすることにより、符号化歪が大きい場合に復号音の品質劣化が少ない音源モードを選択することが可能となり、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質を改善することができるという効果が得られる。

## 【 0 2 0 2 】

## 【発明の効果】

以上のように、この発明によれば、複数の音源モードのそれぞれについて、入力音声から求められる符号化対象信号を符号化して、その時の符号化歪と、固定の閾値あるいは符号化対象信号の信号パワーに基づいて決定した閾値との比較結果に基づいて音源モードを選択し、それを用いて入力音声をフレーム毎に符号化するように構成したので、符号化歪が大きい場合でも復号音の品質劣化が少ない音源モードの選択が可能となって、よりよい音質を与える音源モードの選択を適切に行うことができ、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質、つまり音質が改善できるという効果がある。

## 【 0 2 0 3 】

この発明によれば、複数の音源モードのそれぞれについて、入力音声から求まる符号化対象信号の符号化を行って、その時の符号化歪の相互比較結果より音源モードの1つを選択し、その音源モードに対応する符号化歪と、固定の閾値あるいは符号化対象信号の信号パワーに基づいて決定した閾値とを比較して、その比較結果に基づいた音源モードの置換を行い、それを用いて入力音声をフレーム毎に符号化するように構成したので、符号化歪が大きい場合でも復号音の品質劣化が少ない音源モードの選択が可能となって、よりよい音質を与える音源モードの選択を適切に行うことができ、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質（音質）が改善できるという効果がある。

## 【 0 2 0 4 】

この発明によれば、符号化歪が閾値を上回る場合、比較結果が得られている音源モードの選択を抑制するように構成したので、符号化歪が大きい場合に、復号音の品質劣化が少ない音源モードが選択しやすくなるため、よりよい音質を与える音源モードを適切に選択できるため、得られた音声符号を復号化した復号音の主観的な品質が改善できるという効果がある。

## 【 0 2 0 5 】

この発明によれば、閾値を音源モード毎に用意するように構成したので、各音源モード毎に復号音品質の劣化を引き起こすことを検知する閾値を適切に調整することが可能となって、よりよい音質を与える音源モードを適切に選択できるため、得られた音声符号を復号化して得られる復号音の主観的な品質が改善できるという効果がある。

## 【 0 2 0 6 】

この発明によれば、符号化歪と閾値との比較を、あらかじめ定めた音源モードについて行い、符号化歪が閾値を上回る場合にはその符号化歪を閾値の値に置換して、全音源モードの符号化歪中の最小の符号化歪に対応した音源モードを選択するように構成したので、符号化歪が大きい場合に、符号化歪が置換された音源モードが選択されやすくなって、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善できるという効果がある。

## 【 0 2 0 7 】

この発明によれば、選択した音源モードに対応する符号化歪を選択して閾値と比較し、それが閾値を上回る場合には、あらかじめ定めた音源モードを選択するように構成したので、符号化歪が大きい場合に、復号音の品質劣化が少ない音源モードを強制的に選択することが可能となり、よりよい音質を与える音源モードを適切に選択でき、得られた音声符号を復号化して得られる復号音の主観的な品質が改善できるという効果がある。

## 【 0 2 0 8 】

この発明によれば、閾値を、入力音声または符号化対象信号に対して所定の歪

率となるように設定するように構成したので、符号化時の歪率が所定以上である場合に、復号音の品質劣化が少ない音源モードを選択することが可能となり、よりよい音質を与える音源モードの選択が適切に行え、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善できるという効果がある。

## 【0209】

この発明によれば、入力音声または符号化対象信号を分析して音声様態の判定を行い、所定の判定結果となったときに限り、符号化歪と閾値との比較結果を使用せずに音源モードの選択を行うように構成したので、符号化歪が大きくても復号音の品質劣化を起こしにくい入力音声に対しては、従来の場合と同様の音源モード選択が行われ、一層丁寧な音源モード選択が可能となって、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善できるという効果がある。

## 【0210】

この発明によれば、入力音声または符号化対象信号の分析による音声様態の判定結果に基づいて決定した閾値を用いて比較を行うように構成したので、音声様態に応じて適切に設定された閾値を用いて音源モードの選択を行うことが可能となり、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質が改善できるという効果が得られる。

## 【0211】

この発明によれば、音声様態の判定において、少なくとも音声の立ち上がりであるか否かを判定するように構成したので、符号化歪が大きくなりがちな音声の立ち上がり区間と、それ以外の区間とで、符号化歪に基づく音源モード選択の制御を変えることが可能となるため、音声の立ち上がりでの劣化がなく、それ以外での音源モード選択を改善することができて、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善でき、また、音声の立ち上がり区間では、破裂音などのように雑音的な音源よりパルス的な音源の方が適するケースもあり、符号化歪が大きくても、特定の音源モードを優先的に選択するような制御では劣化を引き起こすことがあるが、音声の立ち上がりの判定によってそれを回避できるなどの効果がある。

## 【 0 2 1 2 】

この発明によれば、非雑音的な音源を生成する音源モードと、雑音的な音源を生成する音源モードとによって、複数の音源モードを形成するように構成したので、符号化歪が大きい場合に、雑音的な音源を生成する音源モードの選択がしやすくなるため、非雑音的な音源を生成する音源モードを選択することによる劣化を回避し、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善できるという効果がある。

## 【 0 2 1 3 】

この発明によれば、非雑音的な音源符号語を使用する音源モードと、雑音的な音源符号語を使用する音源モードとによって、複数の音源モードを形成するように構成したので、符号化歪が大きい場合に、雑音的な音源符号語を使用する音源モードを選択しやすくなるため、非雑音的な音源符号語を使用する音源モードを選択して劣化することを回避し、得られた音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善できるという効果がある。

## 【図面の簡単な説明】

【図 1】 この発明の実施の形態 1 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 2】 この発明の実施の形態 2 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 3】 この発明の実施の形態 3 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 4】 この発明の実施の形態 4 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 5】 この発明の実施の形態 5 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 6】 この発明の実施の形態 6 による音声符号化方法を適用した音声符号化装置の構成を示すブロック図である。

【図 7】 音声符号を音声復号化装置で復号化した復号音の主観的な品質が改善されることを説明するための波形図である。

【図 8】 従来の音声符号化装置の構成の一例を示すブロック図である。

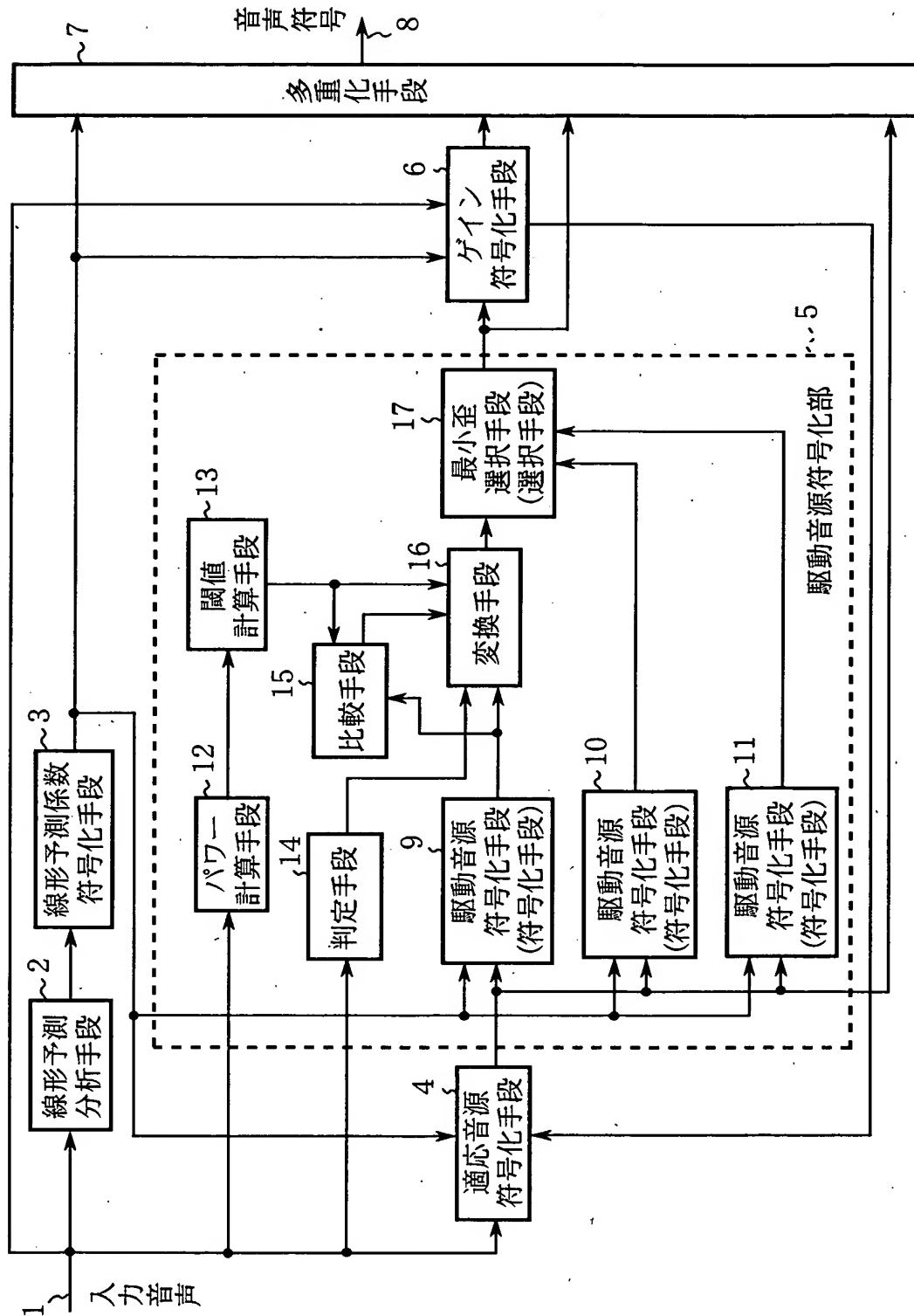
【図 9】 従来の音声符号化装置の構成の他の例を示すブロック図である。

【符号の説明】

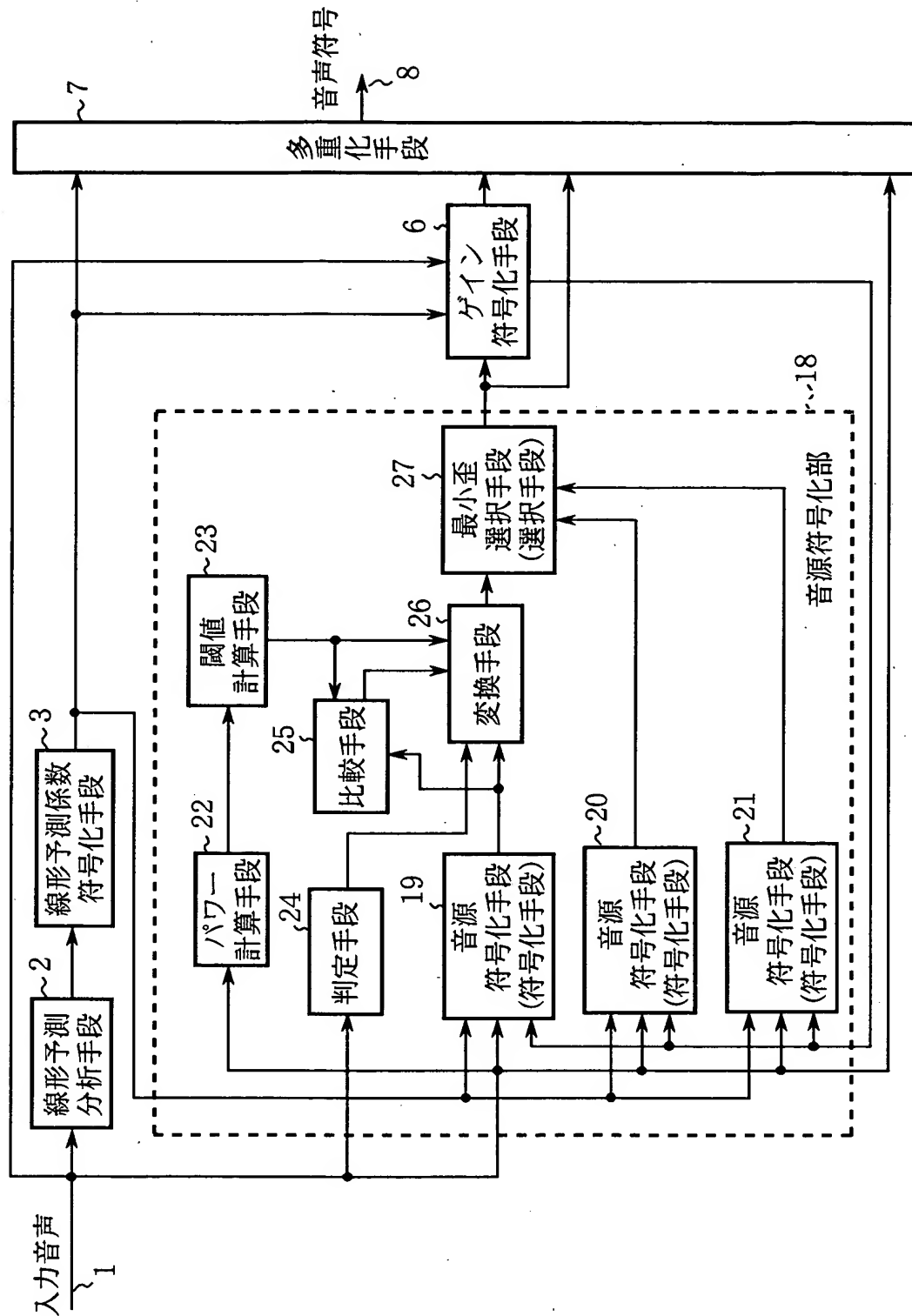
1 入力音声、2 線形予測分析手段、3 線形予測係数符号化手段、4 適応音源符号化手段、5 駆動音源符号化部、6 ゲイン符号化手段、7 多重化手段、8 音声符号、9, 10, 11 駆動音源符号化手段（符号化手段）、12 パワー計算手段、13 閾値計算手段、14 判定手段、15 比較手段、16 変換手段、17 最小歪選択手段（選択手段）、18 音源符号化部、19, 20, 21 音源符号化手段（符号化手段）、22 パワー計算手段、23 閾値計算手段、24 判定手段、25 比較手段、26 変換手段、27 最小歪選択手段（選択手段）、28 駆動音源符号化部、29 閾値計算手段、30, 32 比較手段、31, 33 補正手段（変換手段）、34 駆動音源符号化部、35 最小歪選択手段（選択手段）、36 比較手段、37 置換手段、38 駆動音源符号化部、39 判定手段、40 閾値計算手段、41 変換手段、42 駆動音源符号化部、43, 45 駆動音源符号帳、44, 46 駆動音源符号化手段（符号化手段）。

【書類名】 図面

【図 1】

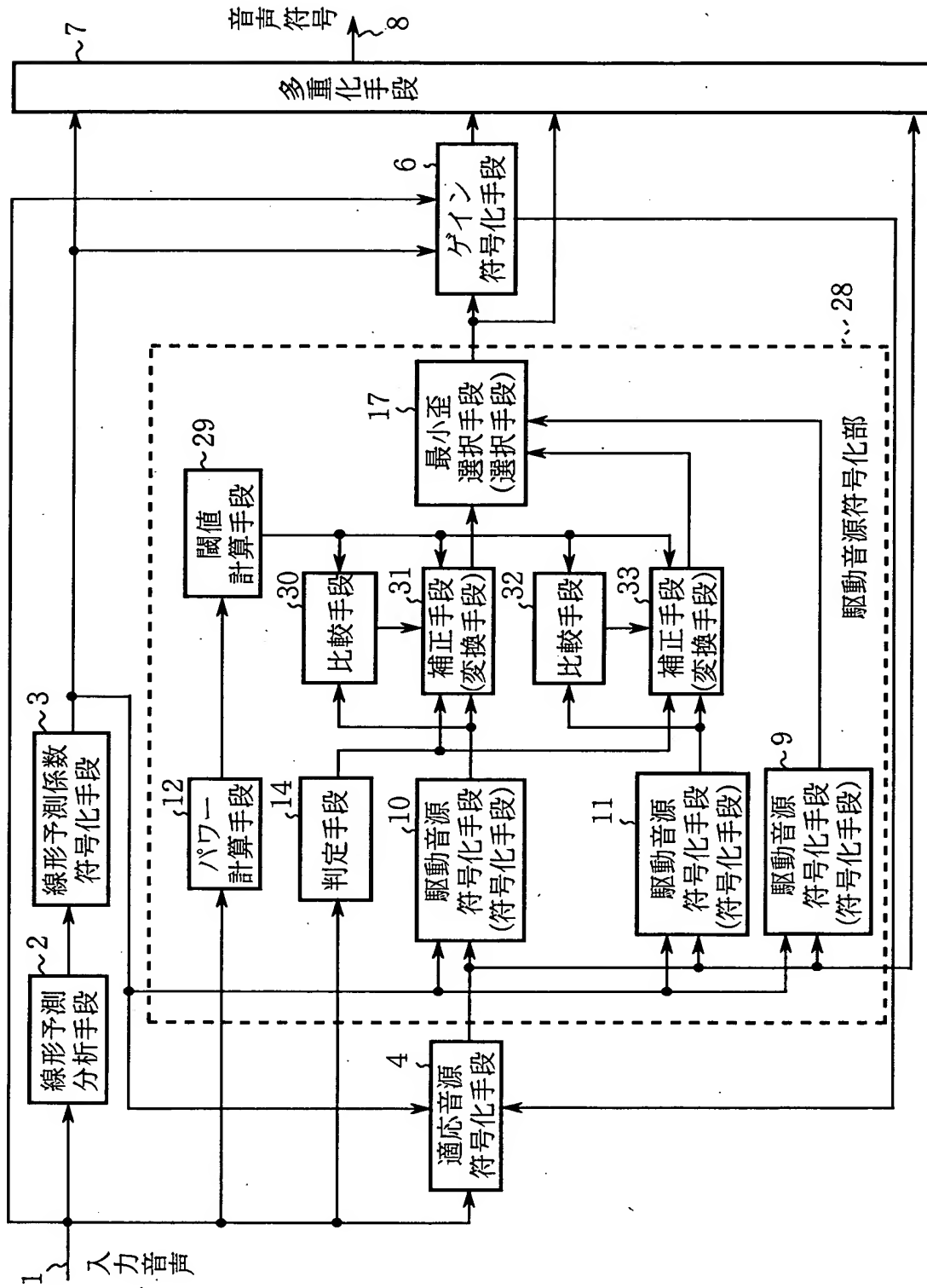


【図 2】

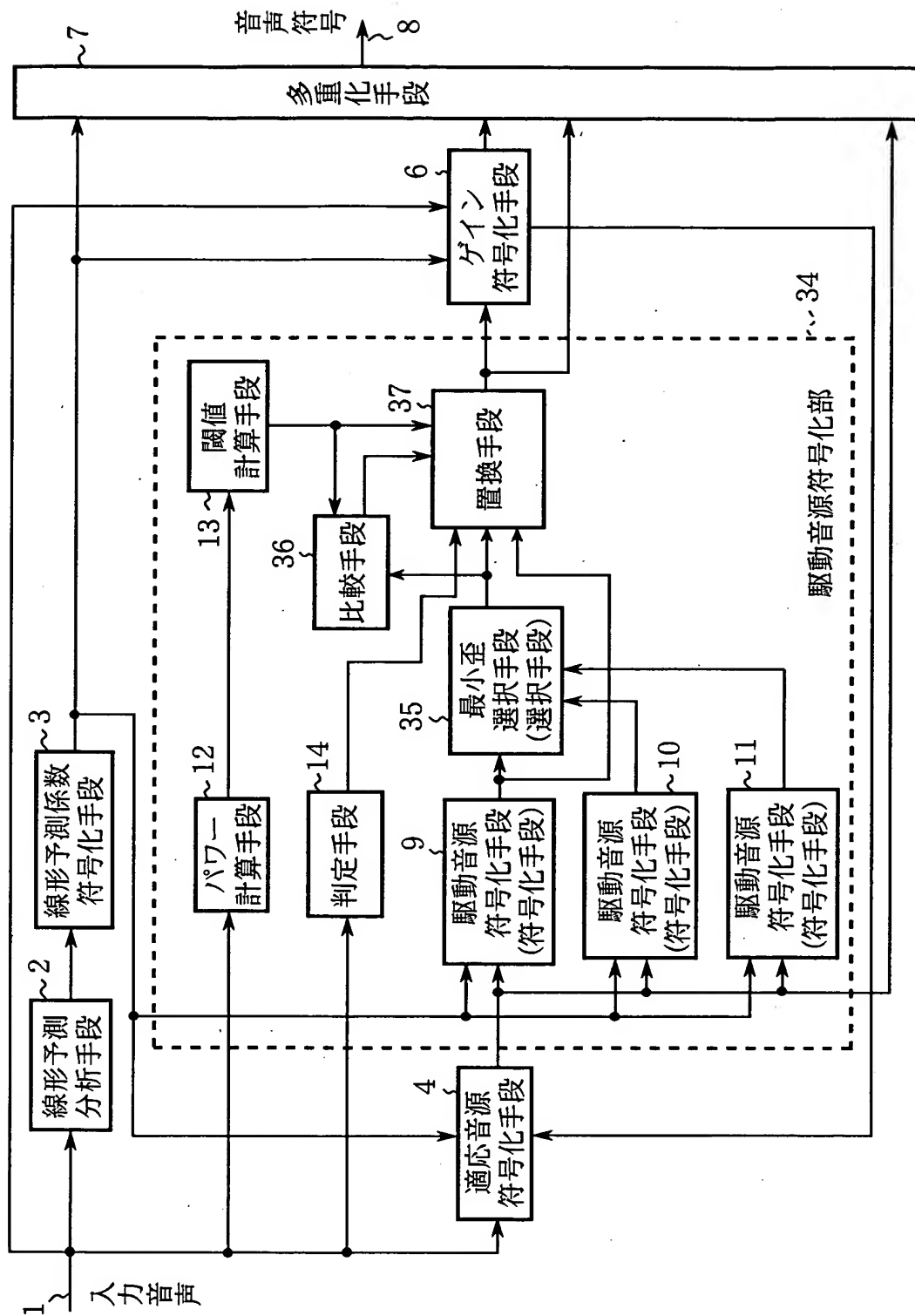




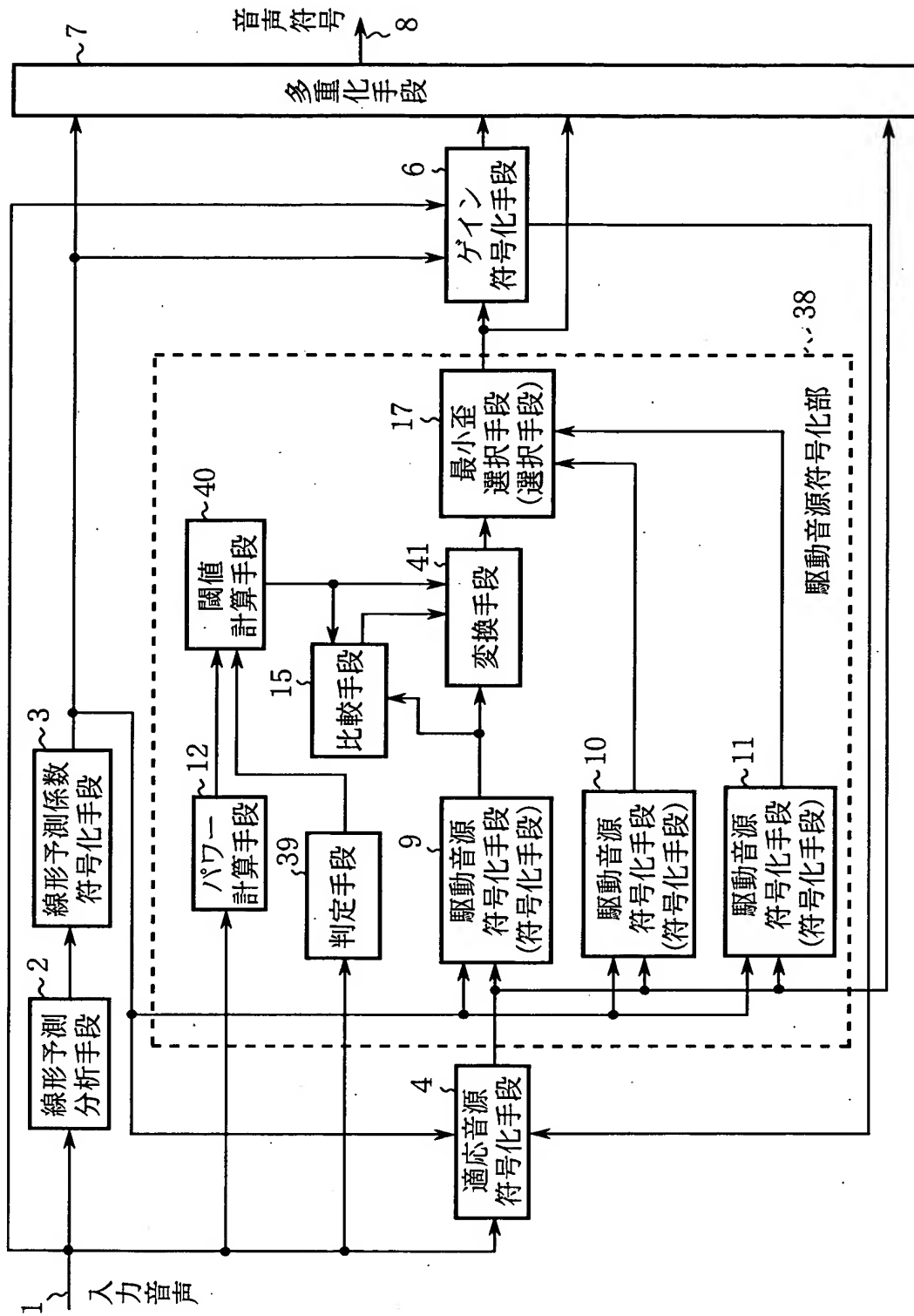
【図 3】



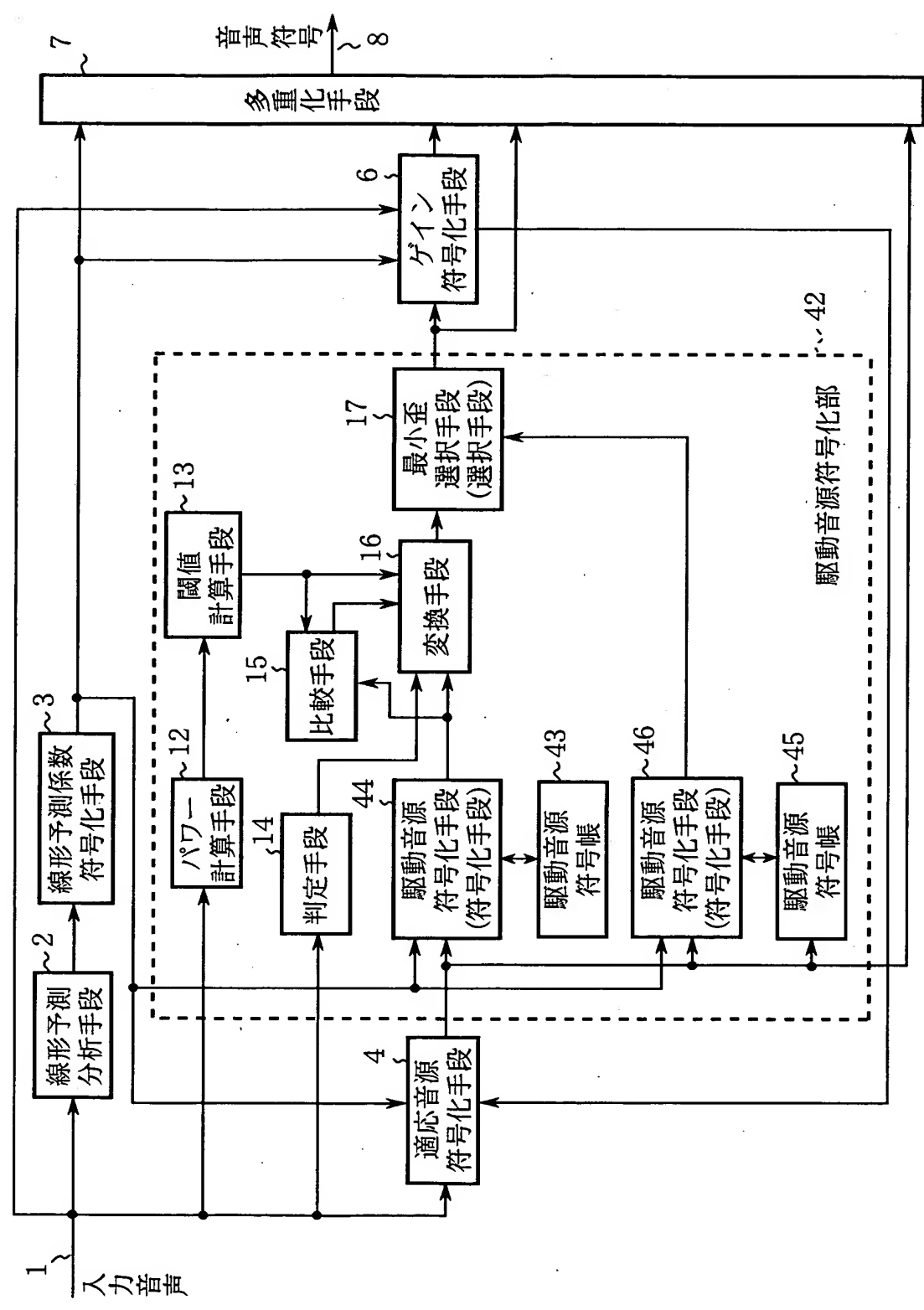
【図 4】



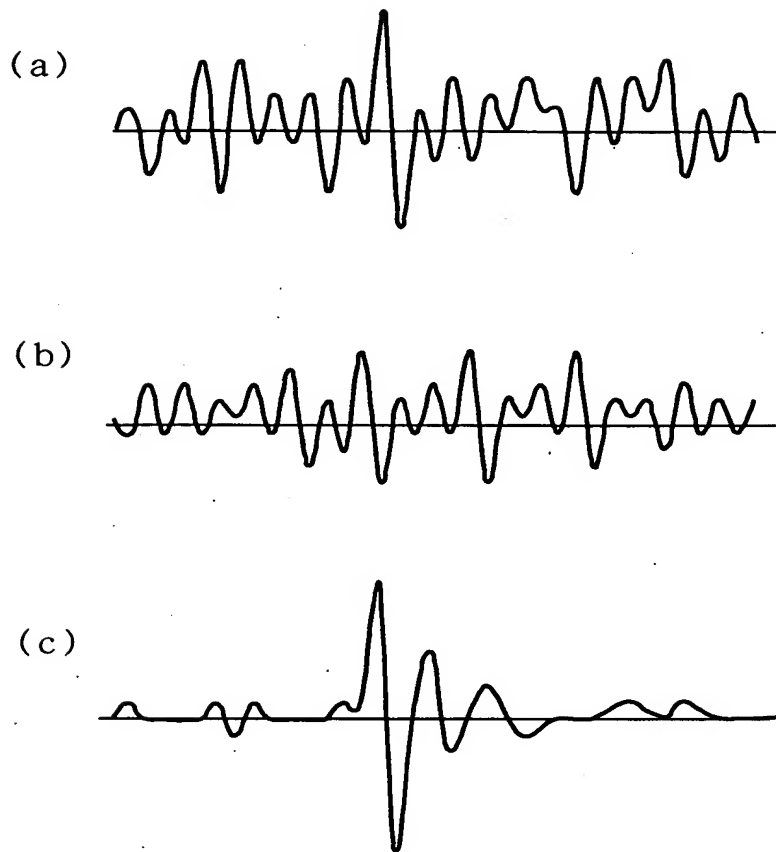
【図 5】



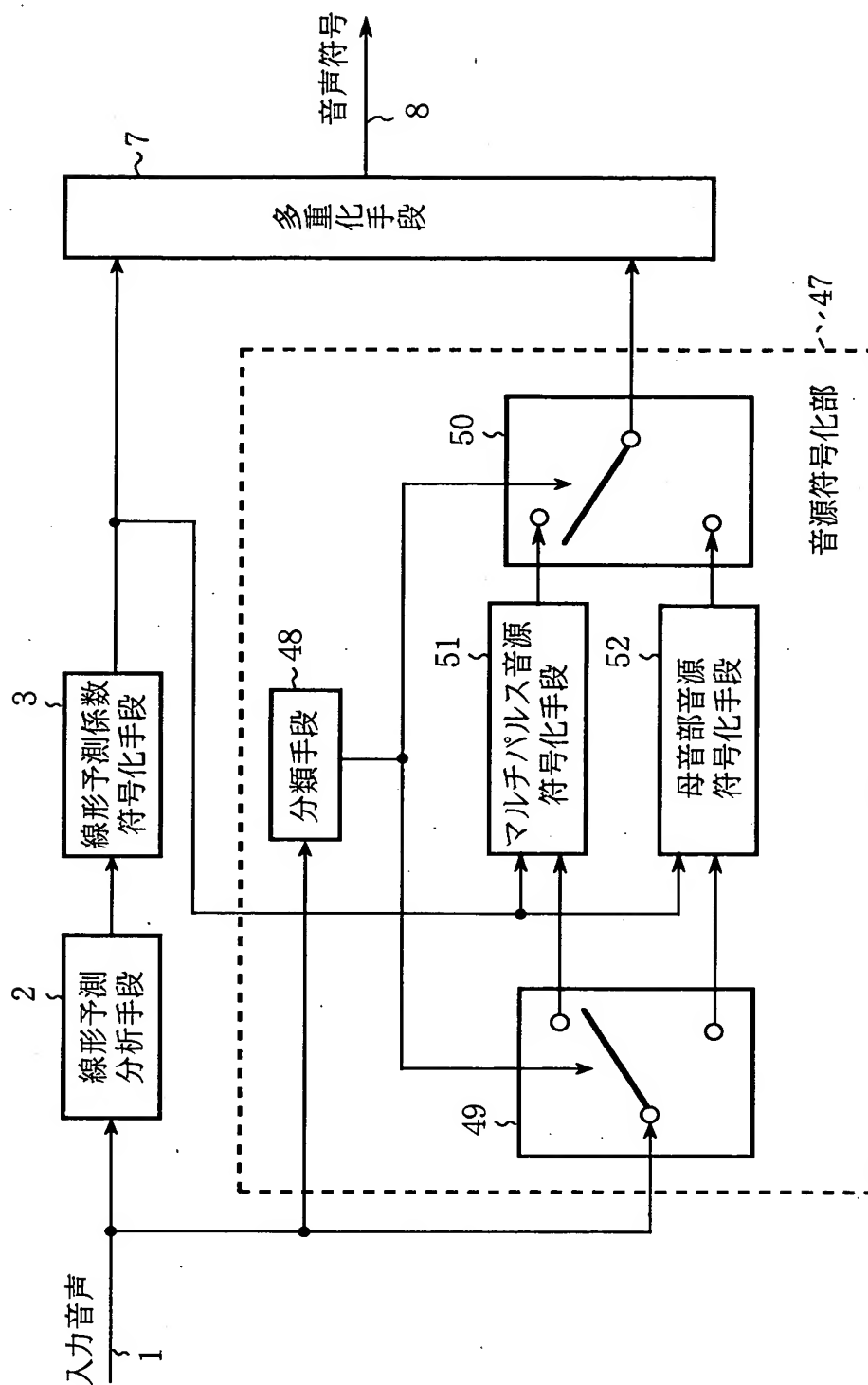
【図 6】



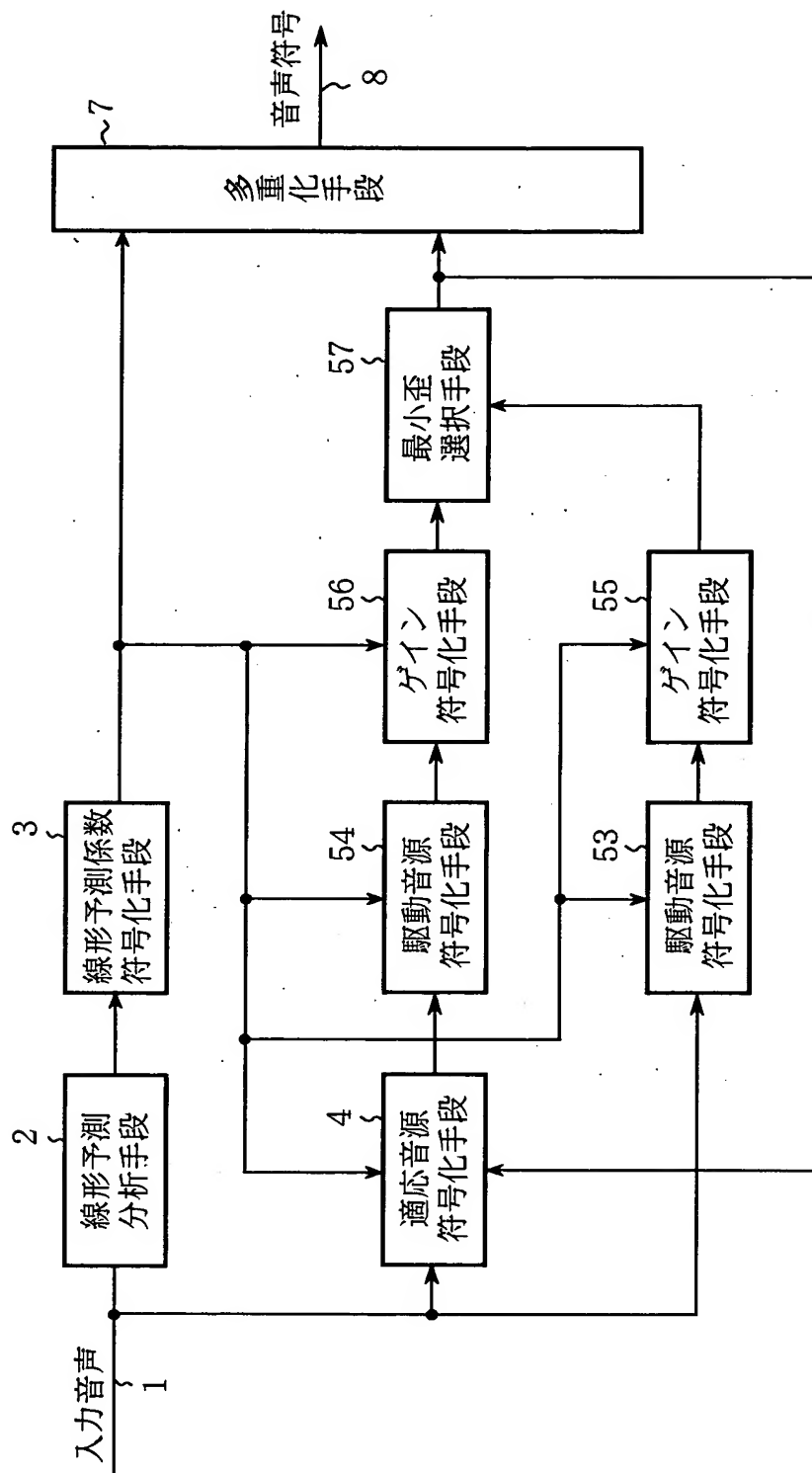
【図7】



【図 8】



【図 9】



【書類名】            要約書

【要約】

【課題】    よりよい音質を与える音源を適切に選択でき、得られた音声符号を音声復号化装置で復号化して得られる復号音の主観的な品質を改善することができる音声符号化方法および音声符号化装置を得る。

【解決手段】    各音源モード毎に、入力音声1から求まる符号化対象信号の符号化を行って、その時の符号化歪を出力する駆動音源符号化手段9～11、符号化歪と、固定の閾値または符号化対象信号の信号パワーに基づいて決定した閾値との比較を行う比較手段15、および、符号化歪と、比較手段の比較結果に基づいて、音源モードの選択を行う最小歪選択手段17を備えたものである。

【選択図】            図1



出 願 人 履 歴 情 報

識別番号 [000006013]

1. 変更年月日	1990年 8月24日
[変更理由]	新規登録
住 所	東京都千代田区丸の内2丁目2番3号
氏 名	三菱電機株式会社